LABORATORY MEDICINE PRACTICE GUIDELINES
Laboratory Support for the Diagnosis and Monitoring of Thyroid Disease

Table of Contents

Editors:
Laurence M. Demers, Ph.D., F.A.C.B.
Carole A. Spencer Ph.D., F.A.C.B.

Guidelines Committee:
The preparation of this revised monograph was achieved with the expert input of the editors, members of the guidelines committee, experts who submitted manuscripts for each section and many expert reviewers, who are listed in Appendix A. The material in this monograph represents the opinions of the editors and does not represent the official position of the National Academy of Clinical Biochemistry or any of the co–sponsoring organizations.
The National Academy of Clinical Biochemistry is the official academy of the American Association of Clinical Chemistry.

We gratefully acknowledge the following individuals who contributed the original manuscripts upon which this monograph is based:

Zubair Baloch, M.D., Ph.D.,
University of Philadelphia Medical Center,
Philadelphia, PA, USA

Pierre Carayon, M.D., D.Sc
U555 INSERM and Department of Biochemistry & Molecular Biology,
University of the Medeiterranea Medical School, Marseille, France

Bernard Conte-Devolx, M.D. Ph.D
U555 INSERM and Department of Endocrinology,
University of the Medeiterranea Medical School, Marseille, France

Ulla Feldt Rasmussen, M.D.
Department of Medicine, National University Hospital,
Copenhagen, Denmark

Jean-François Henry M.D.
U555 INSERM and Department of Endocrine Surgery,
University of the Medeiterranea Medical School, Marseille, France

Virginia LiVolsi, M.D.
University of Philadelphia Medical Center,
Philadelphia, PA, USA

Patricia Niccoli-Sire, M.D.
U555 INSERM and Departments of Endocrinology and Surgery
University of the Mediterranea Medical School, Marseille, France

Rhys John, Ph.D., F.R.C.Path,
University Hospital of Wales, Cardiff, Wales, UK

Jean Ruf, M.D.
U555 INSERM and Department of Biochemistry & Molecular Biology,
University of the Medeiterranea Medical School, Marseille, France

Peter PA Smyth, M.D.
National University of Ireland, Dublin, Ireland

Carole A. Spencer, Ph.D., F.A.C.B.
University of Southern California, Los Angeles, California, USA

Jim R. Stockigt, M.D., F.R.A.C.P., F.R.C.P.A.,
Ewen Downie Metabolic Unit, Alfred Hospital, Melbourne, Victoria, Australia

Section 1. Forward and Introduction

Physicians need quality laboratory testing support for the accurate diagnosis and cost-effective management of thyroid disorders. On occasion, the clinical suspicion is strong, as in clinically overt hyperthyroidism in a young adult or a in the presence of a rapidly growing thyroid mass. In these instances, laboratory thyroid hormone testing simply confirms the suspicion. However in the majority of patients, thyroid disease symptoms are subtle in presentation so that only biochemical testing or cytopathologic evaluation can detect the disorder. However overt or

obscure a patient's thyroid problem may be, an open collaboration between physicians and clinical laboratory scientists is essential for optimal, cost–effective management of the patient with thyroid disease.

Thyroid dysfunction, especially thyroid insufficiency caused by a deficiency in iodine, is a worldwide problem. Iodine deficiency is not always uniform across a nation. Studies in both Europe and the United States suggest it should be considered more as a "pocket disorder", meaning that it can be more prevalent in some areas of a country compared with others (1-3). The creation of this updated monograph is a collaborative effort between the major thyroid organizations worldwide: the American Association of Clinical Endocrinologists (AACE), the Asia & Oceania Thyroid Association (AOTA), the American Thyroid Association (ATA), the British Thyroid Association (BTA), the European Thyroid Association (ETA) and the Latin American Thyroid Society (LATS). These organizations are the authoritative bodies that spearhead thyroid research and have published standards of care for treating thyroid disease in each region of the world. Because geographic and economic factors impact the clinical use of thyroid tests to some extent, this monograph will focus on the technical aspects of thyroid testing and the performance criteria needed for optimal clinical utility of thyroid tests in an increasingly cost-sensitive global environment. Different thyroid testing strategies are favored by individual clinicians and laboratories around the world (5). This monograph cannot accommodate all these variations in thought and opinion but we hope that readers of this monograph will appreciate this. We believe, however, that most of the commonly performed tests and diagnostic procedures used to diagnose and treat thyroid disorders are described in this text. The monograph is designed to give both clinical laboratory scientists and practicing physicians an overview regarding the current strengths and limitations of those thyroid tests most commonly used in clinical practice. Consensus recommendations are made throughout the monograph. The consensus level is > 95%, unless otherwise indicated. We continue to welcome constructive comments that would improve the monograph for a future revision.

A. Additional Resources

Current clinical guidelines are published in the following references (5-11). In addition, the textbooks “Thyroid” and “The Thyroid and Its Diseases” (www.thyroidmanager.org) are useful references (12,13). A list of symptoms suggesting the presence of thyroid disease together with the ICD-9 codes recommended to Medicare by the American Thyroid Association is available on the ATA website (www.thyroid.org). Clinical practice guidelines may vary, depending on the region of the country. More information can be obtained from each of the thyroid organizations: Asia & Oceania Thyroid Association (AOTA = www -dnm.kuhp.kyoto-u.ac.jp/AOTA); American Thyroid Association (ATA = www.thyroid.org); European Thyroid Association (ETA =www.eurothyroid.com) and Latin American Thyroid Society (LATS = www.lats.org).

B. Historical Perspective

Over the past forty years, improvements in the sensitivity and specificity of biochemical thyroid tests, as well as the development of fine needle aspiration biopsy (FNA) and improved cytological techniques, have dramatically impacted clinical strategies for detecting and treating thyroid disorders. In the 1950s, only one serum-based thyroid test was available - an indirect estimate of the total (free + protein-bound) thyroxine (T4) concentration, using the protein bound iodine (PBI) technique. Today, urine iodine concentrations are measured directly by dry or wet-ash techniques and are used to estimate dietary iodine intake. The development of radioimmunoassays (RIA) in the early 1970s and more recently, non-isotopic immunometric assay (IMA) methods have progressively improved the specificity and sensitivity of thyroid hormone testing. Currently, serum-based tests are available for measuring the concentration of both the total (TT4 and TT3) and free (FT4 and FT3) thyroid hormones in the circulation (14,15). In addition, measurements of the thyroid hormone binding plasma proteins, Thyroxine Binding globulin (TBG), Transthyretin (TTR)/Prealbumin (TBPA) and Albumin are available (16). Improvements in the sensitivity of assays to measure the pituitary thyroid stimulating hormone, thyrotropin (TSH) now allow TSH to be used for detecting both hyper- and hypothyroidism. Further, measurement of the thyroid gland precursor protein, Thyroglobulin (Tg) as well as the measurement of Calcitonin (CT) in serum have become important tumor markers for managing patients with differentiated and medullary thyroid carcinomas, respectively. The recognition that autoimmunity is a major cause of thyroid dysfunction has led to the development of more sensitive and specific tests for autoantibodies to thyroid peroxidase (TPOAb), thyroglobulin (TgAb) and the TSH receptor (TRAb). Current thyroid tests are

usually performed on serum by either manual or automated methods that employ specific antibodies (17). Methodology continues to evolve as performance standards are established and new technology and instrumentation are developed.

Section 2. Pre-Analytic Factors

Thyroxine (T4) is the principal hormone secreted by the thyroid gland. All T4 in the circulation is derived from thyroidal secretion. In contrast, only about 20% of circulating triiodothyronine (T3) is directly secreted by the thyroid gland. Most of the T3 circulating in blood is produced enzymatically in nonthyroidal tissues from 5'-monodeiodination of T4. In fact, T4 appears to function as a pro-hormone for the production of the more biologically active form of thyroid hormone, T3. In the circulation, most of the (~99.98%) T4 is bound to specific plasma proteins, thyroxine-binding globulin (TBG) (60-75%), TTR/TBPA (prealbumin/transthyretin) (15-30%) and albumin (~10%) (12,16). T3 is ten-fold more weakly protein-bound than T4, with ~99.7% protein-bound, primarily to TBG (12). Protein-bound thyroid hormones do not enter cells and are thus considered to be biologically inert and function as storage reservoirs for circulating thyroid hormone. In contrast, the minute free hormone fractions readily enter cells by specific membrane transport mechanisms to exert their biological effects. In the pituitary, the negative feedback of thyroid hormone on TSH secretion is mediated primarily by T3 that is produced at the site from the free T4 entering the thyrotroph cells.

Fortunately, most pre-analytic variables have little effect on serum TSH measurements - the most common thyroid test used initially to assess for thyroid status in ambulatory patients. Pre-analytic variables and interfering substances present in specimens may influence the binding of thyroid hormones to plasma proteins and thus decrease the diagnostic accuracy of total and free thyroid hormone measurements, more frequently than serum TSH (see Table 1). As discussed in [Section 2B(b) and 3B(d)viii] both FT4 and TSH values may be diagnostically misleading in the hospitalized setting of severe nonthyroidal illness (NTI). Indeed, euthyroid patients frequently have abnormal serum TSH and/or total and free thyroid hormone concentrations as a result of NTI, or secondary to medications that might interfere with hormone secretion or synthesis. When there is a strong suspicion that one of these variables might affect test results, consulting advice from the expert physician or clinical biochemist is frequently needed.

Table 1. Causes of FT4/TSH Discordance in the Absence of Serious Associated Illness

Misleading Test	Result TSH	Result FT4	Likely Causes	Suggested Action
FT4	↑	N	1. Untreated - mild hypothyroidism 2. L-T4 Rx. - inadequate/noncompliance	1. Measure TPOAb. Confirm TSH after 6 weeks 2. Increase L-T4 dose/council compliance
	↓	N or ↓	1. Mild (subclinical) hyperthyroidism 2. L-T3 treatment	1. ? Autonomous functioning goiter. 2. Measure FT3 to rule out T3-toxicosis. 3. Overtreatment with T3-containing preparation
	N	↑	1. Common during L-T4 treatment. 2. Abnormal Binding proteins (i.e. FDH) 3. Competitor drugs or IV Heparin 4. HAMA or rheumatoid factor interference	1. Expect higher FT4 in L-T4 Rx. hypo. 2, 3 & 4. Check FT4 by alternate FT4 method ideally one using physical separation i.e. equilibrium dialysis or ultrafiltration
	N	↓	1. Binding-protein competitors /other drugs 2. Pregnancy & Hypoalbuminemia 3. Overdosing with anti-thyroid drugs 4. Euthyroid Pregnancy	1. Check FT4 by method using minimal dilution 2. Check FT4 by an albumin-insensitive method 3. Reduce dose and recheck in 3 weeks 4. If L-T4 Rx., monitor TSH & FT4 each trimester
TSH	↑	N	1. Dysequilibrium (first 6-8 weeks of L-T4 Rx. for 1° hypothyroidism) 2. HAMA & other interferences	1. Recheck TSH before adjusting L-T4 dose. High TSH may persist for months after Rx. severe hypo. 2. Check TSH (new spec.)by alternate method
	↓	N	1. Dysequilibrium (first 2-3 months post Rx. for hyperthyroidism) 2. Medications, i.e. glucocorticoids, dopamine	1. Use FT4 and FT3 during early Rx. of hyper to monitor thyroid status. TSH may take months to normalize after starting Rx. for severe hyperthyroidism
	N or ↑	↑	1. Heterophilic Antibody (HAMA) 2. TSH-secreting pituitary adenoma 3. During first 6 months of amiodarone Rx. 4. Thyroid Hormone Resistance (rare)	1. Check TSH in new specimen by alternate method 2. TRH-stim or thyroid hormone suppression test 3. TSH alpha subunit/Pituitary Imaging 4. Genetic studies of thyroid hormone receptor
	N	↓	1. Central hypothyroidism associated with secretion of a TSH with normal immuno-activity but reduced bioactivity	1. ? Other signs of pituitary deficiency 2. ? Blunted (< 2-fold/≤+4mU/L) TRH response

- *Abnormal TSH + normal FT4 discordances are common. These usually indicate mild (subclinical) hypo- or hyperthyroidism when the pituitary is normal and thyroid status stable.*
- *Abnormal FT4 + normal TSH discordances usually indicate an interference with the FT4 measurement, often secondary to abnormal T4-binding proteins.*

In addition to basic physiologic variability, individual patient variables such as genetic abnormalities in thyroid binding proteins or severe nonthyroidal illness (NTI) may impact the sensitivity and specificity of thyroid test. In addition, iatrogenic factors such as thyroid and nonthyroidal medications such as glucocorticoids or beta-blockers; and specimen variables, including autoantibodies to thyroid hormones and Tg as well as heterophilic antibodies (HAMA) can affect the diagnostic accuracy resulting in test result misinterpretation. Table 2 lists the pre-analytic factors to consider when interpreting thyroid tests.

A. *Physiologic Variables*

For practical purposes, variables such as age, gender, race, season, phase of menstrual cycle, cigarette smoking, exercise, fasting or phlebotomy-induced stasis have minor effects on the reference intervals for thyroid tests in ambulatory adults (4). Since the differences in these physiological variables are less than the method-to-method differences encountered in clinical practice they are considered inconsequential.

Table 2

A. Physiologic Variables
- TSH/free T4 relationship
- Age
- Pregnancy
- Biologic variation

B. Pathologic Variables
- Thyroid gland dysfunction
- Hepatic or renal dysfunction
- Medications
- Systemic illnesses

C. Specimen-related Variables
- Interfering factors

2.1 General Guidelines for Laboratories & Physicians

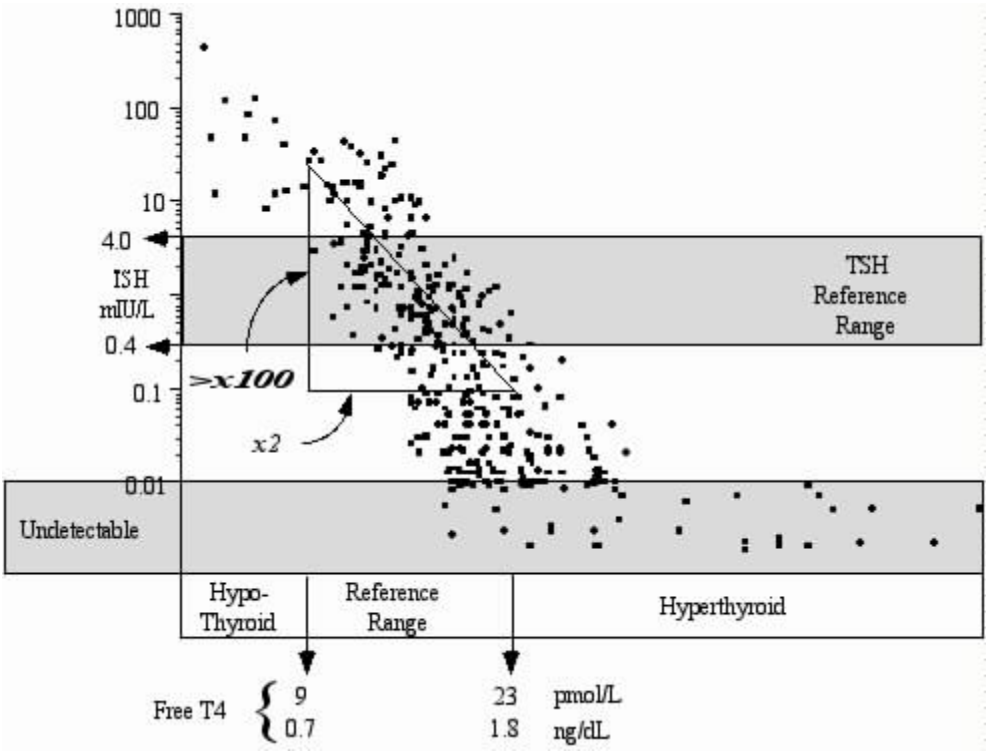
- Laboratories should store (at 4-8°C) all serum specimens used for thyroid testing for at least one week after the results have been reported to allow physicians time to order additional tests.
- Specimens from differentiated thyroid cancer patients (DTC) sent for serum Thyroglobulin (Tg) measurement should be archived (at -20°C) for a minimum of six months.

(a) <u>The serum TSH/ FT4 relationship</u>

An understanding of the normal relationship between serum levels of free T4 (FT4) and TSH is essential when interpreting thyroid tests. Needless to say, an intact hypothalamic-pituitary axis is a prerequisite if TSH measurements are to be used to determine primary thyroid dysfunction (18). A number of clinical conditions and pharmaceutical agents disrupt the FT4/TSH relationship. As shown in Table 1, it is more common to encounter misleading FT4 tests than misleading serum TSH measurements.

When hypothalamic-pituitary function is normal, a log/linear inverse relationship between serum TSH and free T4 concentrations is produced by negative feedback inhibition of pituitary TSH secretion by thyroid hormones. Thus, thyroid function can be determined either directly, by measuring the primary thyroid gland product, T4 (preferably as free T4) or indirectly, by assessing the TSH level which inversely reflects thyroid hormone concentrations sensed by the pituitary. It follows that high TSH and low FT4 is characteristic of hypothyroidism and low TSH and high FT4 is characteristic of hyperthyroidism). In fact, now that the sensitivity and specificity of TSH assays have improved, it has become apparent that the indirect approach (serum TSH measurement) offers better sensitivity for detecting thyroid dysfunction than does FT4 testing.

Figure 1. The TSH / Free T4 Relationship



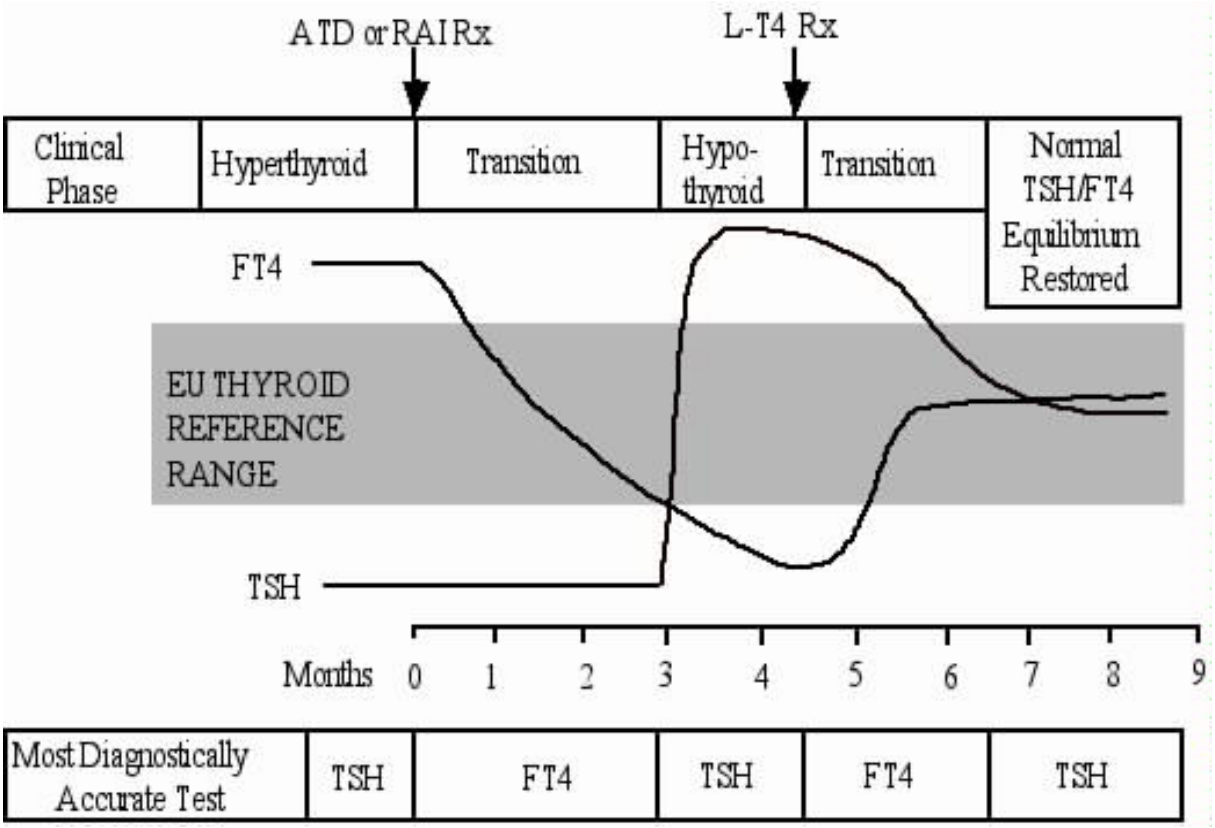
There are two reasons for using a TSH-centered strategy for ambulatory patients:

- Firstly, as shown in Figure 1, serum TSH and FT4 concentrations exhibit an inverse log/linear relationship such that small alterations in FT4 will produce a much larger response in serum TSH (19).

- Secondly, twin studies show that each individual has a genetically determined FT4 set-point (20). This dictates that any mild FT4 excess or deficiency will be sensed by the pituitary, relative to that individual's set-point, and cause an amplified, inverse response in TSH secretion. Thus, in the early stages of developing thyroid dysfunction, a serum TSH abnormality will precede the development of an abnormal FT4 because of the exponential TSH response to even subtle FT4 changes. At the early stage of thyroid disease development, serum TSH becomes abnormal while FT4 remains within population reference limits. This is because population limits are broad having been constructed from a cohort of individuals with different FT4 set-points.

Currently, measurement of the serum TSH concentration is our most reliable indicator of thyroid status at the tissue level. For example, studies of mild (subclinical) thyroid hormone excess or deficiency (abnormal TSH/normal range FT4 and FT3) find abnormalities in markers of thyroid hormone action in a variety of tissues (heart, brain, bone, liver and kidney). These abnormalities typically reverse when treatment is initiated to normalize serum TSH (21-24).

Figure 2. Lag in Pituitary TSH Reset during Unstable Thyroid Status Following Rx.



It is important to recognize the clinical situations when serum TSH or FT4 levels may be diagnostically misleading (see Table 1). These include:

- Abnormalities in hypothalamic or pituitary function, including TSH-producing pituitary tumors.
- Unstable thyroid status, such as occurs in the early phase of treating primary hypothyroidism or changing L-T4 dose (6 to 8 weeks). Additionally, following the treatment for hyperthyroidism it takes time (2 to 3 months or more) for the serum TSH concentration to re-equilibrate to the new thyroid hormone status (Figure 2).
- Drugs that influence pituitary TSH secretion (i.e. dopamine and glucocorticoids) or thyroid hormone binding to plasma proteins.

2.2 General Guidelines for Thyroid Testing of Ambulatory Patients

- ***Patients with <u>stable</u> thyroid status:*** When thyroid status is stable and hypothalamic-pituitary function is intact, serum TSH measurement is more sensitive than a free T4 (FT4) test for detecting mild (subclinical) thyroid hormone excess or deficiency. The superior diagnostic sensitivity of serum TSH reflects the log/linear relationship between TSH and FT4 and the exquisite sensitivity of pituitary to sense free T4 abnormalities relative to an individual's genetically determined free T4 set-point.
- ***Patients with <u>unstable</u> thyroid status:*** When thyroid status is unstable (i.e. during the first 2-3 months of treatment for hypo- or hyperthyroidism) serum FT4 measurement more reliably indicates the new thyroid status than does TSH. Patients with chronic, severe hypothyroidism may develop pituitary thyrotroph hyperplasia which can mimic a pituitary adenoma, but which resolves after several months of L-T4 replacement therapy. Both TSH and FT4 should be used for monitoring hypothyroid patients suspected of intermittent or non-compliance with L-T4 replacement therapy. Such patients may exhibit discordant serum TSH and FT4 values (high TSH/high FT4) because of persistent disequilibrium in the FT4 / TSH relationship.

(b) Age

(i) Adults.

Despite studies showing minor differences between older and younger subjects, adult age-adjusted reference ranges for thyroid hormones and TSH are unnecessary (4,25,26). With respect to euthyroid elderly individuals, the TSH mean value increases each decade as do the prevalences for both low and high serum TSH concentrations compared with younger individuals (4,27,28). Despite the greater serum TSH variability seen in older individuals, there appears to be no justification for using a widened or age-adjusted reference ranges (25,26). This conservative approach is justified by reports that mildly suppressed or elevated serum TSH is associated with increased cardiovascular morbidity and mortality (29,30).

(ii) Neonates, Infants and Children.

In children, the hypothalamic/pituitary/thyroid axis undergoes progressive maturation and modulation. Specifically, there is a continuous decrease in the TSH/FT4 ratio from the time of mid-gestation until after the completion of puberty (31-36). This maturation process dictates the use of age-specific reference limits. However, there are significant differences between FT4 and TSH measurements made by different methods [see Sections 3B and 3C]. Since most manufacturers have not independently established age-specific reference limits, these limits can be calculated for different assays by adjusting the upper and lower limits of the adult range by the ratio between child and adult values, such as indicated in Table 3.

Lower serum total and free FT3 concentration (measured by most methods) are seen with pregnancy, during the neonatal period, in the elderly and during caloric deprivation (15). Conversely, in children higher TSH concentrations are typically seen (37). Further, higher total and free FT3 concentrations are typically seen in euthyroid children. This suggests that the upper T3 limit for young patients (less than 20 years of age) should be established as a gradient: between 6.7 pmol/L (44 ng/dL) for adults, up to 8.3 pmol/L (54 ng/dL) for children under three years of age (38).

Table 3. Relative TSH and FT4 Reference Ranges during Gestation and Childhood

Age	TSH Child/ Adult Ratio	TSH Ranges mU/L	FT4 Child/ Adult Ratio	FT4 Ranges pmol/L (ng/dL)
Midgestation Fetus	2.41	0.7-11	0.2	2-4 (0.15-0.34)
LBW cord serum	4.49	1.3-20	0.8	8-17 (0.64-1.4)
Term infants	4.28	1.3-19	1.0	10-22 (0.8-1.9)
3 days	3.66	1.1-17	2.3	22-49 (1.8-4.1)
10 weeks	2.13	0.6-10	1.0	9-21 (0.8-1.7)
14 months	1.40	0.4-7.0	0.8	8-17 (0.6-1.4)
5 years	1.20	0.4-6.0	0.9	9-20 (0.8-1.7)

14 years	0.97	0.3-5.0	0.8	8-17 (0.6-1.4)
Adult	**1.00**	**0.3-4.0**	**1.0**	**9-22 (0.8-1.8)**

2.3 Guidelines for Thyroid Testing of Children

- The hypothalamic-pituitary-thyroid axis matures throughout infancy until the completion of puberty.
- Both TSH and FT4 concentrations are higher in children, especially in the first week of life and throughout the first year. Failure to recognize this could lead to a missed and/or under-treatment of congenital hypothyroidism. Age-related normal reference limits should be used for all thyroid tests (see Table 3).

(c) Pregnancy

During pregnancy states, estrogen excess progressively increases mean TBG concentrations that plateau at 2 to 3 times pre-pregnancy levels by 20 weeks of gestation (39,40). This rise in TBG results in a shift in the TT4 and TT3 reference ranges to approximately 1.5 times non-pregnant levels by 16 weeks of gestation (41-43). These changes result in a fall in serum TSH during the first trimester period, such that subnormal serum TSH may be seen in approximately 20 % of normal pregancies (39,40,44). This decrease in TSH is attributed to the thyroid stimulating activity of human chorionic gonadotropin (hCG) which has structural homology with pituitary TSH (45-47). The peak rise in hCG and the nadir in serum TSH occur together at about 10-12 weeks of gestation. In approximately 10 % of such cases (i.e. 2 % of all pregnancies) the increase in free T4 reaches supranormal values and, when prolonged, may lead to a syndrome entitled "gestational transient thyrotoxicosis" (GTT) that is characterized by more or less pronounced symptoms and signs of thyrotoxicosis pregnancy (45,47).This condition is frequently associated with hyperemesis in the first trimester of (48,49).

2.4 Guidelines for Thyroid Testing of Pregnant Patients

- Mounting evidence suggests that mild hypothyroidism during early pregnancy has a detrimental effect on outcome (fetal loss and infant IQ). This suggests that pre-pregnancy or first trimester screening for thyroid dysfunction with serum TSH would be efficacious. Initiation of levothyroxine (L-T4) therapy should be considered if the serum TSH level is above the reference range for that stage of pregnancy.
- A high serum TPOAb concentration during the first trimester is a risk factor for post-partum thyroiditis.
- Serum TSH should be used to assess thyroid status during each trimester when pregnant patients are taking L-T4 therapy.
- Trimester-specific reference intervals should be used when reporting thyroid test values for pregnant patients.
- In general, after 16 weeks of pregnancy, both TT4 and TT3 reference ranges are 1.5 fold higher than the non-pregnant ranges.
- FT3 and FT4 reference ranges in pregnancy are method-dependent and should be established independently for each method.
- Serum thyroglobulin (Tg) concentrations rise during normal pregnancy and also when patients with DTC have thyroid tissue present.

The fall in TSH during the first trimester of pregnancy is associated with a modest increase in FT4 (39,40,44). Thereafter, in the second and third trimesters of pregnancy there is now consensus that serum FT4 and FT3 levels decrease to approximately 20 to 40 percent below the normal mean, a decrease in free hormones that is further amplified when the iodine nutrition status of the mother is restricted or deficient (39,40,44). In some patients, FT4 may fall below the lower reference limit for non-pregnant patients (50-52,44). The frequency of subnormal FT4

concentrations in this setting is method-dependent (50,53-57). Patients receiving L-T4 replacement therapy who become pregnant may require an increased dose to maintain normal serum TSH (58,59). Thyroid status should be checked with TSH + FT4 during each trimester. L-T4 dose should be adjusted to maintain normal TSH and FT4 concentrations. Serum Tg concentrations typically rise during normal pregnancy (39). In patients with differentiated thyroid carcinomas (DTC) with thyroid tissue still present typically show a two-fold rise in serum Tg with a return to baseline by 6 to 8 weeks postpartum.

Decreased availability of maternal thyroxine may be a critical factor for fetal neurologic development in the early stages of gestation, before the fetal thyroid gland becomes active. Several recent studies report both increased fetal loss as well as IQ deficits for infants born to mothers with either undiagnosed hypothyroidism, low range FT4 or TPOAb positivity (60-63). However, one study suggests that early identification and treatment of mild (subclinical) hypothyroidism may prevent the long-term effects of low thyroid hormone levels on the psychomotor and auditory systems of the neonate (64).

(d) Biologic Variation

The intra-individual serum levels of the thyroid hormones as well as their precursor protein, thyroglobulin (Tg) are quite stable over a 1 to 4 year period (Table 4) (65). All thyroid analytes display a greater inter-individual variability compared to intra-individual variability (Table 4) (65,66). The stability of intra-individual serum T4 concentrations reflects the long half-life (7 days) of thyroxine and the genetic imposed free T4 setpoint. The stability of intra-individual T3 concentrations reflects autoregulation of the rate of T4 to T3 conversion (67). Inter-individual variability is especially high for serum Tg concentrations, because individuals have differences in thyroid mass, TSH status and may have conditions associated with thyroid injury (i.e. thyroiditis) – all conditions that influence serum Tg concentrations. Serum TSH levels also display high variability, both within and between individuals. This primarily reflects the short half-life of TSH (~60 minutes) together with its circadian and diurnal variations, the latter peaking during the night and reaching a nadir sometime between 1000 to 1600 hrs (68,69). The amplitude of the diurnal TSH variability across a 24-hour period is approximately 2-fold (68,69). However, since this magnitude of change lies within the normal TSH reference interval for the population as a whole (~0.3 to 4.0 mIU/L) it does not affect the utility of TSH for diagnosing thyroid dysfunction. Furthermore, TSH is typically measured during the day in the outpatient setting when the magnitude of the TSH diurnal variability is smaller.

B. Pathologic Variables

(a) Medications

Medications can cause both in vivo and in vitro effects on thyroid tests. This may lead to misinterpretation of laboratory data and inappropriate diagnoses, unnecessary further testing and escalating health care costs (70,71). **In Vivo Effects:** In general, the serum TSH concentration is affected less by medications than thyroid hormone levels (Table 1). For example, Estrogen-induced TBG elevations raise serum TT4 levels but do not affect the serum TSH concentration, because TSH is controlled by the FT4 level that is independent of binding-protein effects. Glucocorticoids in large doses can lower the serum T3 level and inhibit TSH secretion (72,73). Dopamine also inhibits TSH secretion and may even mask the raised TSH level of primary hypothyroidism in sick hospitalized patients (74). Propranolol is sometimes used to treat manifestations of thyrotoxicosis and has an inhibitory effect on T4 to T3 conversion. Propranolol given to non-thyroid disease individuals for their hypertension can cause an elevation in TSH as a result of the impaired T4 to T3 conversion (75). Iodide, contained in solutions for sterilizing the skin and radioopaque dyes and contrast media used in CAG and CT-scans, can cause both hyper and hypothyroidism (76). In addition, the iodine-containing cardiac drug Amiodarone has complex effects on thyroid gland function that may induce either hypothyroidism or hyperthyroidism in susceptible patients with positive TPOAb (77-81). Amiodarone-induced hyperthyroidism has two classifications. Type 1 appears to be induced in abnormal thyroid glands by the excess of iodine contained in the drug. A combination of thionamide drugs and potassium perchlorate has often been used in such cases. Type II appears to result from a destructive thyroiditis that is often treated with prednisone and thionamide drugs. Some studies report elevated IL-6 levels in Type II. The measurement of serum T3 levels is necessary to confirm the diagnosis of Amiodarone-induced hyperthyroidism. L-

T4-treated patients who take Amiodarone may have disproportionately elevated serum TSH concentrations relative to the serum FT4 level (81). Lithium can cause hypo- or hyperthyroidism in as many as 10% of lithium-treated patients, especially those with positive TPOAb (82-84).

2.5 Guidelines for Patients taking Amiodarone Medication

Amiodarone therapy can induce the development of hypo- or hyperthyroidism in 14-18% patients with apparently normal thyroid glands or with preexisting abnormalities.

 Pretreatment – thorough physical thyroid examination together with baseline thyroid tests (TSH, TPOAb, FT4 and FT3) is recommended before beginning Amiodarone therapy.
 First 6 months. Abnormal tests may occur in the first six months after initiating therapy. TSH may be discordant with thyroid hormone levels (high TSH/highT4/low T3). TSH usually normalizes with long-term therapy if patients remain euthyroid.
 Long-term follow-up. Monitor thyroid status every 6 months with TSH + FT4 + FT3. Serum TSH is the most reliable indicator of thyroid status during therapy.
 Hypothyroidism. Preexisting Hashimotos' thyroiditis and/or TPOAb-positivity is a risk factor for developing hypothyroidism at any time during therapy.
 Hyperthyroidism. T3 (total and free) usually remains low during therapy. An inappropriately normal or high T3 is suspicious for hyperthyroidism. Two types of amiodarone-induced hyperthyroidism may develop during therapy, although mixed forms are frequently seen.

- *Type 1* = excess iodine-induced. Recommended treatment = simultaneous administration of thionamides and potassium perchlorate and in some cases, thyroidectomy. Seen more often in areas of low iodine intake. However, in iodine-sufficient areas, radioiodine uptakes may be low precluding radioiodine as a therapeutic option.
  - *Type 1a*: most commonly seen in patients with nodular goiters.
  - *Type 1b:* less commonly seen during developing Graves' disease.
- *Type II* = amiodarone-induced destructive thyroiditiss. Recommended treatment = glucocorticoids and or beta-blockers. Radioiodine uptakes are typically low or suppressed. Type II is most commonly seen in iodine-sufficient areas.

In Vitro Effects: Some therapeutic and diagnostic agents may be present in serum specimens in sufficient concentrations to produce in vitro interference with some thyroid test methods. The therapeutic agents themselves (i.e. Phenytoin, Carbamazepine or Furosemide/Frusemide) may competitively inhibit thyroid hormone binding to serum proteins in the specimen, resulting in a reduction in serum TT4 and FT4 values. Such effects on FT4 measurements are usually an artifact of the commercial assay designs, since they are not seen with methods such as equilibrium dialysis and ultrafiltration which employ physical isolation of the FT4 fraction (85,86). Alternatively, following I.V. Heparin administration, in-vitro stimulation of lipoprotein lipase can liberate FFA, which inhibits T4 binding to serum proteins (87). In certain pathologic conditions such as uremia, abnormal serum constituents such as indole acetic acid may accumulate and interfere with thyroid hormone binding (88). Thyroid test methods employing fluorescent signals may be sensitive to the presence of fluorophore-related therapeutic or diagnostic agents in the specimen (89)

(b) Severe Nonthyroidal Illness (NTI)

Patients who are seriously ill often have abnormalities in their thyroid tests but usually do not have thyroid dysfunction (90,91). These abnormalities, which are seen with both acute and chronic critical illnesses, are thought to arise from a central maladjusted inhibition of hypothalamic releasing hormones, including TRH (92,93). The terms "nonthyroidal illness" or NTI, as well as "euthyroid sick" are often used to describe this subset of patients (94). The spectrum of changes in thyroid tests relates both to the severity and stage of illness, as well as to technical factors affecting the methods and in some cases the medications used in these patients.

2.6 Guidelines for Testing of Hospitalized Patients with Non Thyroidal Illness (NTI)

- Acute or chronic non-thyroidal illness has complex effects on thyroid function tests and whenever possible diagnostic testing should be deferred until the illness has resolved, except when the patient's history or clinical features suggest the presence of thyroid dysfunction.
- Physicians should recognize that some thyroid tests are inherently un-interpretable in severely sick patients, or patients receiving multiple medications.
- All estimates of free or total T4 measurements in NTI should be interpreted in conjunction with the serum TSH concentration that in the absence of dopamine administration, is the more reliable test. The combined assessment (T4 +TSH) is the most reliable way of distinguishing true primary thyroid dysfunction from the transient abnormalities of NTI.
- Reverse T3 (rT3) testing is rarely helpful in the hospital setting, because paradoxically normal or low values can result from impaired renal function. Further, the test is not readily available.
- An abnormal FT4 test in the setting of serious somatic disease is unreliable, since the FT4 methods used by clinical laboratories lack diagnostic specificity for evaluating sick patients.
- An abnormal FT4 result in a hospitalized patient should be confirmed by a reflex TT4 measurement. If both TT4 and FT4 are abnormal (in the same direction) a thyroid condition may be present. When TT4 and FT4 are discordant, the FT4 abnormality is unlikely due to thyroid dysfunction and more likely a result of the illness, medications or an artifact of the test.
- TT4 abnormalities should be assessed relative to the severity of illness, since the low TT4 state of NTI is typically only seen in severely sick patients with high mortality. A low TT4 concentration in a patient not in intensive care is suspicious for hypothyroidism.
- A raised total or free T3 is a useful indicator of hyperthyroidism in a hospitalized patient, but a normal or low T3 does not rule it out.

It is well established that both FT4 and TSH measurements have reduced specificity for detecting thyroid dysfunction when patients have severe non-thyroidal illnessess, as compared with ambulatory patients (19,95,96). It is generally recommended that thyroid function testing of hospitalized patients be limited to those patients with clinical symptoms or a history of thyroid dysfunction (96). The reasons for the reduced specificity of thyroid testing in this setting are multifactorial. Many such patients receive medications such as dopaminergics, glucocorticoids, furosemide or heparin that either inhibit pituitary TSH secretion or directly or indirectly inhibit T4 binding to proteins, as discussed above. Further, there is circumstantial evidence that binding protein affinities are reduced, possibly by the presence of endogenous inhibitors that circulate in some pathologic conditions (57,97-100).

Most hospitalized patients have low serum TT3 and FT3 concentrations, as measured by most methods (14,101). As the severity of the illness increases, serum TT4 typically fall because of a disruption of binding protein affinities, possibly caused by T4-binding inhibitors in the circulation (94,102,103). It should be noted that subnormal TT4 values only develop when a critical illness (usually sepsis) is present. Such patients are usually in an ICU setting. If TT4 is low and yet the patient is not profoundly sick, central hypothyroidism secondary to pituitary or hypothalamic deficiency should be considered.

FT4 and FT3 estimate values are method dependent and may be either spuriously high or low, depending on the methodologic principles underlying the test. For example, FT4 tests are unreliable if they are sensitive to the release of FFA following IV heparin infusion or are sensitive to dilution artifacts (87,97,101,102,104,105). FT4 methods that physically separate free from protein-bound hormone such as equilibrium dialysis and ultrafiltration usually generate normal or elevated values for critically ill patients [see Section 2B(b) and 3B(d)viii] (97,106).

Serum TSH levels remain within normal limits in the majority of NTI patients, provided that no dopamine or glucocorticoid therapy is administered (90,96). However, in acute NTI there may be a mild, transient fall in serum TSH to the 0.02-0.2 mIU/L range, followed by a rebound to mildly elevated values during recovery (107). Thus it is important to use a TSH assay that has a functional sensitivity at or below 0.02 mIU/L in the hospitalized setting. Without this level of sensitivity it is not possible to reliably determine the degree of TSH suppression which helps discriminate sick hyperthyroid patients with profoundly low serum TSH values (< 0.02 mIU/L) from patients with merely a mild transient suppression caused by a NTI. Minor elevations in TSH are also less diagnostic for

hypothyroidism in the hospitalized setting. Sick hypothyroid patients typically exhibit the combination of low FT4 and elevated TSH (>20 mIU/L) (95).

It is clear that the diagnosis and treatment of thyroid dysfunction in the presence of a severe NTI is not simple, and is best done with the help of an endocrine specialist. The use of a TT4 measurement +/- a THBR estimate of binding protein capacity_ may be more diagnostically helpful than using current FT4 immunoassay tests that have variable diagnostic accuracy for assessing sick patients. Although the diagnostic specificity of TSH is reduced in the presence of somatic illnesses, a detectable serum TSH value (using an appropriately sensitive assay) usually rules out significant thyroid dysfunction, provided that hypothalamic-pituitary function is intact and the patient is not receiving medications_____ that affect pituitary TSH secretion_____. In general, it is best to avoid routine thyroid testing in acutely ill patients. Empiric treatment of the low TT4 state of NTI has not improved outcome and is still considered experimental (108-110).

C. Specimen Variables

(a) Stability

A few studies have examined the effects of maintaining blood samples at ambient on the serum concentrations of the total and free thyroid hormones, TSH and thyroglobulin levels (111). In general, these studies suggest that thyroid hormones are relatively stable whether stored at room temperature, refrigerated or frozen. Some studies have reported that T4 in serum is stable for months when stored at 4°C and years when frozen at –10°C (112,113). TSH and TT4 in dried whole blood spots used to screen for neonatal hypothyroidism are also stable for months when stored with a desiccant. TSH in serum has been reported to be slightly more stable than T4 (114). It is important to note however, as discussed above, that non-frozen specimens from patients receiving heparin are prone to in-vitro generation of FFA that can spuriously elevate FT4 when measured by some methods (87).

(b) Serum Constituents

Hemolysis, lipemia, and hyperbilirubinemia do not produce significant interference in immunoassays, in general. However, free fatty acids can displace T4 from serum binding proteins, which may partly explain the low TT4 values often seen in NTI (104).

2.7 Guidelines for Investigating Discordant Thyroid Test Results in Ambulatory Patients

Discordant thyroid test results can result from technical interference or rare clinical conditions

- ***Technical Interferences****:* Technical interference can sometimes be detected by measuring the specimen with a different manufacturer's method, since the magnitude of most interference is method-dependent. Alternatively, non-linearity in dilutions of the specimen may indicate a technical interference with TT4, TT3 or TSH measurements. It is not recommended to dilute most FT4 and FT3 tests used by clinical laboratories, because such tests are binding protein dependent and do not usually give linear dilution responses.

- ***Rare Clinical Conditions****:* Unexpectedly abnormal or discordant thyroid test values may be seen with some rare, but clinically significant conditions such as central hypothyroidism, TSH-secreting pituitary tumors, thyroid hormone resistance, or the presence of heterophilic antibodies (HAMA) or thyroid hormone (T4 and/or T3) autoantibodies.

(c) Heterophilic Antibodies (HAMA)

These antibodies may be encountered in patient sera. They affect immunometric assay (IMA) methods more than competitive immunoassays by forming a bridge between the capture and signal antibodies, thereby creating a false signal, resulting in a falsely high or low value depending on the methodology see [Section 3B(f)] (115,116). The inappropriate result may not necessarily be abnormal, but merely be inappropriately normal. Since antibodies cross

the placenta, HAMA has the potential to influence the neonatal screening result. Manufacturers are currently employing various approaches to deal with the HAMA issue with varying degrees of success including the use of chimeric antibody combinations and blocking agents to neutralize the effects of HAMA on their methods. (117)

(d) Sample Collection and Processing

Serum is the preferred specimen over EDTA- or heparinized-plasma by most manufacturers. For optimal results and maximum serum yield, it is recommended that whole blood samples be allowed to clot for at least 30 minutes before centrifugation and separation. Storage of serum at –20°C is recommended if the assay is to be delayed for more than 24 hours. Collection of serum in barrier gel tubes does not affect the results of most TSH and thyroid hormone tests.

(e) Thyroid Test Performance Standards

The performance of a laboratory test can be evaluated both biologically and analytically. Table 4 shows the biological variation in various thyroid serum analytes, expressed as inter-individual variability and intra-individual variability, across different periods of time (65,66,118,119).

Table 4. Intra-individual and Inter-individual Variability in Serum Thyroid Tests

Serum Analyte	Time span	%CV*	%CV**
TT4 /FT4	1 week	3.5	10.8
	6 weeks	5.3	13.0
	4 years	7.3	14.8
TT3 /FT3	1 week	8.7	18.0
	6 weeks	5.6	14.8
	4 years	15.1	19.2
Thyrotropin (TSH)	1 week	19.3	19.7
	6 weeks	20.6	53.3
	4 years	21.2	45.0
Thyroglobulin (Tg)	1 week	4.4	12.6
	6 weeks	8.7	66.6
	4 months	14.0	35.0

*intra-individual **inter-individual

Analytical performance is typically assessed in the laboratory by parameters such as:

- Within and between-run precision assessed at different analyte concentrations
- Minimum detection limit (analytical sensitivity) (120,121)
- Functional Sensitivity (20% between-run precision)
- Linearity of measurements made across the reportable range of values
- Recovery of analyte added to the standard matrix
- Normal reference interval (mean +/- 2 standard deviation of values) for a cohort of subjects without disease
- Correlation with a reference method

Although analytical performance parameters are the foundation of most laboratory quality control and quality assurance programs, it is universally accepted that analytical performance goals should be established on biology (intra- and inter-individual variation) as well as clinical need. It has been proposed that total analytical error should

ideally be less than half the within-subject biological coefficient of variation (%CV) (119,122-124).

Table 5. Bias and Precision Targets for Thyroid Tests

Test	TT4	**FT4**	**TT3**	**FT3**	**TSH**	**Tg**
Normal	nmol/L(ug/dL)	pmol/L(ng/mL)	nmol/L(ng/dL)	pmol/L(ng/L)	mU/L	µg/L(ng/mL)
Reference Range	58-160/4.5-12.6	9-23/0.7-1.8	1.2-2.7/80-180	3.5-7.7/23-50	0.3-4.0	3.0-40.0
Intra-individual %CV	5.3	9.5	5.6	7.9	20.6	8.7
Inter-individual %CV	13.0	12.1	14.8	22.5	53.3	66.6
W	3.5	3.8	4.0	6.0	14.3	16.8
X	1.3	2.4	1.4	2.0	5.2	2.2
Y	7.0	7.7	7.9	11.9	28.6	33.6
Z	2.7	4.8	2.8	4.0	10.3	4.4
Clinically Significant Differences in Serial Values	**15 (1.0)**	**6 (0.5)**	**0.6 (40)**	**1.5 (9)**	**0.7**	**1.5**

W= Suggested percentage goal for maximum bias in diagnostic testing
X= Suggested percentage goal for maximum bias in monitoring an individual
Y= Suggested percentage goal for maximum imprecision in diagnostic testing
Z= Suggested percentage goal for maximum imprecision for monitoring an individual

For diagnostic testing, thyroid test results are reported together with a "normal" reference interval that reflects inter-individual variability. This reference range provides the benchmark for case finding. Reference ranges, however, cannot be used to determine whether differences between two consecutive test results, made during monitoring a patient's treatment, constitute a clinically significant change, or merely reflect the technical (between-run imprecision) and biologic (intra-individual variance) variability of the measurement (125). The "normal" reference interval is usually irrelevant during post-operative clinical management using tumor markers, such as Tg (126). Clearly, method bias and precision goals need not be as stringent when a measurement is used for diagnostic testing compared with using serial measurements for patient monitoring. While the "normal" reference interval stated on the typical laboratory report helps the physician to make a primary diagnosis, it does not give relevant information to help the physician assess the significance of changes resulting from treatment.

Table 5 shows bias and precision goals for the principle thyroid tests for both diagnostic testing and monitoring. The values shown are calculated from studies of within- and inter-individual precision estimates from various studies and are based on well-established concepts (65,66,118,124,127,128). This table also calculates the magnitude of change in consecutive results (approximating mean normal analyte concentrations) that constitute a clinically significant difference in each test (66). These benchmarks should help the physician gauge the clinical significance of changes seen during serial monitoring the treatment of patients with thyroid disorders.

2.8 Guidelines for Interpreting Thyroid Test Results

- For diagnostic testing (case-finding) thyroid test results are typically reported together with a "normal" reference interval that reflects inter-individual variability.
- The "normal" reference interval does not indicate the magnitude of difference between test results that constitutes a clinically significant change.
- Studies of analytical variability together with inter-individual and intra-individual estimates of biological variability, allow calculation of differences in thyroid test values that can be considered clinically significant when monitoring a patient's for response to therapy:

TT4 = 15 (1.0) nmol/L (ug/dL)

FT4 =	**6 (0.5) pmol/L (ng/mL)**
TT3 =	**0.6 (40) nmol/L (ng/dL)**
FT3 =	**1.5 (9) pmol/L (ng/L)**
TSH =	**0.7 mU/L**
Tg =	**1.5 µg/L (ng/mL)**

Section 3: Thyroid Tests for the Clinical Biochemist and Physician

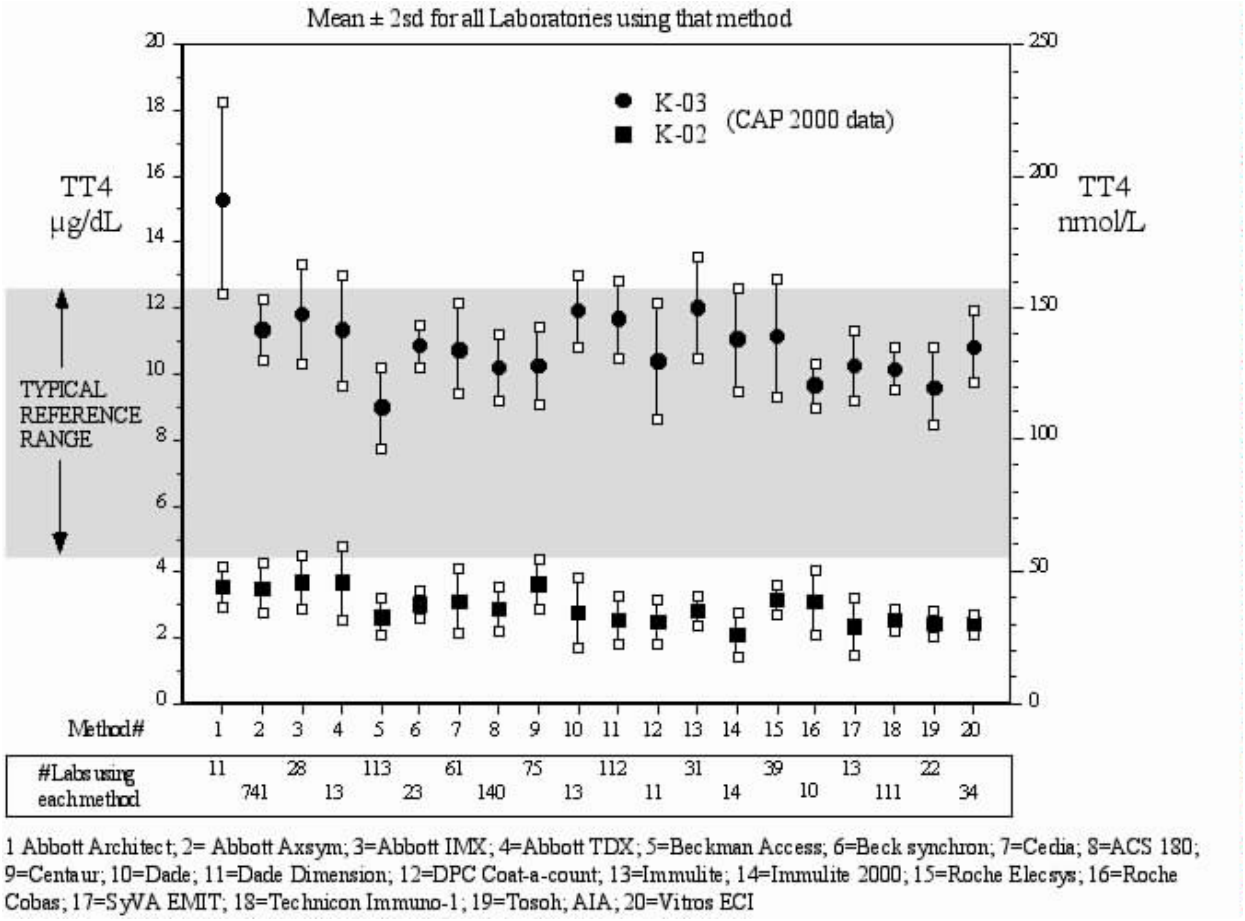
A. Total Thyroxine (TT4) and Total Triiodothyronine (TT3) Methods

Technically, it has been easier to develop methods to measure total (free + protein-bound) thyroid hormone concentrations than tests for the minute free hormone moieties. This is because total hormone concentrations (TT4 and TT3) are measured at nanomolar concentrations whereas free hormone concentrations (FT4 and FT3) are measured at the picomolar level.

(a) Methods for measuring Total Thyroid Hormone Concentrations in Serum

Serum total T4 and total T3 methods (TT4 and TT3) have evolved through a variety of technologies over the past four decades. The PBI tests of the 1950s that estimated the TT4 concentration as "protein-bound iodine" was replaced in the 1960s first by competitive protein binding methods and later in the 1970s by radioimmunoassay (RIA) methods. Currently, serum TT4 and TT3 concentrations are measured either by competitive immunoassay or non-competitive immunometric assay (IMA) methods that use radioactive iodine, enzymes, fluorescence or chemiluminescence molecules as signals. Total hormone assays necessitate the inclusion of an inhibitor (displacing or blocking agent) such as 8-anilino-1-napthalene-sulphonic acid (ANS), or salicylate to release the hormones from binding proteins (129). The displacement of hormone binding from serum proteins by such agents, together with the large sample dilution employed in modern assays, facilitates the binding of hormone to the antibody reagent. The ten-fold lower TT3 concentration, as compared with TT4 in blood, presents both a sensitivity and precision challenge technically, despite the use of a higher specimen volume (130). Although reliable high-range TT3 measurement is critical for diagnosing hyperthyroidism, reliable normal-range measurement is also important for adjusting antithyroid drug dosage and detecting hyperthyroidism in sick hospitalized patients, in whom a paradoxically normal T3 value may indicate hyperthyroidism.

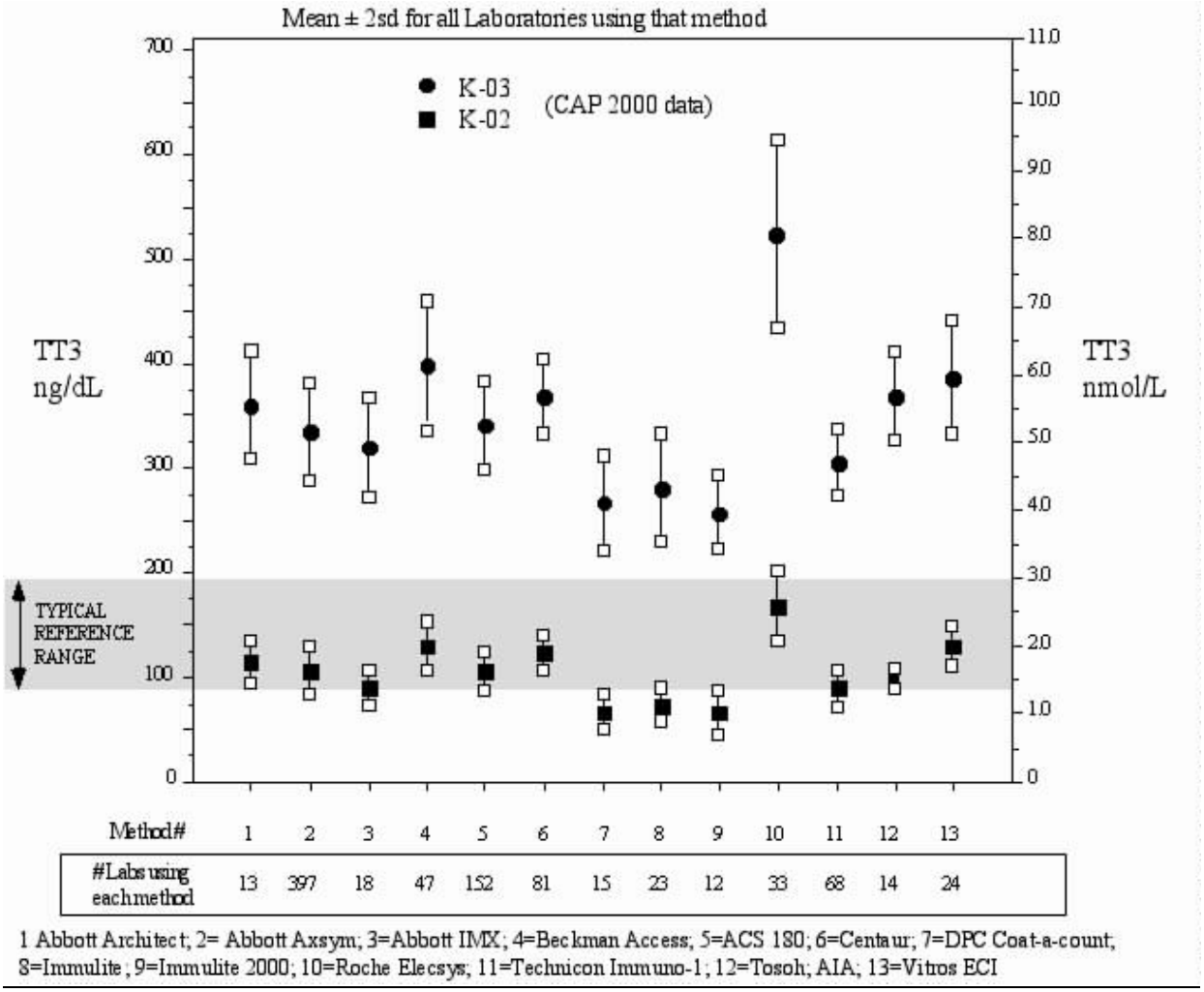
Figure 3. Total T4 Measurement Using Different Methods



As shown in Figures 3 and 4, there are significant between-method biases between different TT4 and TT3 methods. These biases are greater than the maximum allowable bias calculated from the within and between-individual variability of these hormones (Table 5).

A number of factors are likely responsible for these large inter-method biases. Firstly, despite the availability of highly purified preparations of crystalline L-thyroxine and L-triiodothyronine (i.e. from the United States Pharmacopoeia (16201 Twinbrook Parkway, Rockville, MD 20852) no reference method has yet been established (131). Furthermore, the hygroscopic nature of the crystalline preparations can affect the accuracy of gravimetric weighing (132). Secondly, the diluents used to reconstitute L-T4 and L-T3 preparations as calibrators are either modified protein matrices or human serum pools that have been stripped of hormone by various means. In either case, the protein composition of the matrix used for the calibrators is not identical to patient sera. This can result in the protein binding inhibitor reagent (e.g. ANS) releasing different quantities of hormone from calibrator matrix proteins than from the TBG in patient specimens. This may impact the diagnostic accuracy of testing when there are abnormal binding proteins in the specimen, such as in NTI and familial dysalbuminemic hyperthyroxinemia (FDH) (133).

Figure 4. Total T3 Measurement Using Different Methods



(b) Diagnostic Accuracy of Total Hormone Measurements

The diagnostic accuracy of total thyroid hormone measurements would equal that of free hormone if all patients had identical levels of binding proteins (TBG, TTR/TBPA and albumin) with similar affinities for thyroid hormones. Unfortunately, abnormal serum TT4 and TT3 concentrations are more commonly encountered as a result of binding protein abnormalities than for true thyroid dysfunction. Patients with serum TBG abnormalities secondary to pregnancy or estrogen therapy, as well as genetic abnormalities in binding proteins, are frequently encountered in clinical practice (133). Abnormal TBG concentrations and/or affinity for thyroid hormone can distort the relationship between total and free hormone measurements (134). Additionally, some patient sera contain other abnormal binding proteins such as autoantibodies to thyroid hormones that render total hormone measurements diagnostically unreliable (135-137). These binding protein abnormalities compromise the use of TT4 and TT3 measurements as stand-alone thyroid tests. Instead, serum TT4 and TT3 measurements are typically made as part of a two-test panel that includes an assessment of binding protein status, made either directly by TBG immunoassay or by an "uptake" test [Section 3B(b)]. Specifically, a mathematical relationship between the total hormone concentration and the "uptake" result is used as a free hormone "index" (138). Free hormone indices (FT4I and FT3I) have been used as free hormone estimate tests for three decades but are rapidly being replaced by one-test free hormone estimate immunoassays [Section 3B(c)].

A.1 Guidelines for Manufacturers Developing TT4 and TT3 Methods

Methodologic biases should be reduced by:

(b) The development of L-T4 and L-T3 reference preparations and establishing international reference methods.
(c) Ensuring that instruments are not sensitive to differences between human serum and the calibrator matrix.
(d) Ensuring that during the test process, the amount of thyroid hormone released from serum binding proteins is the same as that released in the presence of the calibrator diluent.

(c) Serum TT4 and TT3 Normal Reference Intervals

As shown in Figure 3, serum TT4 values vary to some extent between methods. Reference ranges approximate 58 - 160 nmol/L (4.5-12.6 µg/dL). Likewise, as shown in Figure 4, serum TT3 values are also method dependent, with reference ranges approximating 1.2 - 2.7 nmol/L (80 –180 ng/dL).

A.2 Guidelines for Serum TT4 and TT3 Measurement

- Abnormal serum TT4 and TT3 concentrations are more commonly encountered as a result of binding protein abnormalities and not thyroid dysfunction.
- In ambulatory patients, a free hormone estimate test (index or one-test immunoassay) is usually normal if an abnormal TT4 or TT3 test result is merely due to an abnormality in TBG concentration.
- Free hormone estimate tests differ in diagnostic accuracy when the affinity of TBG and/or the concentration or affinities of other thyroid hormone binding proteins is altered.
- Total T4 assays should remain readily available to evaluate discordant free T4 tests.

B. Free Thyroxine (FT4) and Free Triiodothyronine (FT3) Tests

Thyroxine in blood is more tightly bound to serum proteins than is T3, consequently the free T4 (FT4) bioavailable fraction is less than free T3 (0.02% versus 0.2%, FT4 versus FT3, respectively). Unfortunately, the physical techniques for separating the minute free hormone fractions from the predominant protein-bound moieties are technically demanding, inconvenient to use and relatively expensive for routine clinical laboratory use. These methods that employ physical separation of free from bound hormone (i.e. equilibrium dialysis, ultrafiltration and gel filtration) are typically only available in reference laboratories. The routine clinical laboratories on the other hand uses a variety of free hormone tests that simply estimate the free hormone concentration in the presence of protein-bound hormone. These free hormone estimate tests employ either a two-test strategy to calculate a free hormone "index" [see Section 3B(b)] or a variety of immunoextraction approaches (14,137,140). In reality, despite manufacturers claims, most if not all FT4 and FT3 estimate tests are binding-protein dependent to some extent (52,141). The diagnostic accuracy therefore of many of these free hormone methods is subject to a variety of interference's that can cause misinterpretation or inappropriate abnormal results (Table 1). Such interferences include sensitivity to abnormal binding proteins, in-vivo or in-vitro effects of various drugs [Section 3B(f)], high FFA levels and endogenous or exogenous inhibitors of hormone binding to proteins that are present in certain pathological conditions (57).

B.1 Guideline: Free Hormone Test Nomenclature

- The free hormone methods used by most clinical laboratories (indexes and immunoassays) that do not employ physical separation of bound from free hormone do not measure free hormone concentrations directly! These tests are typically binding protein dependent to some extent and should more appropriately be called **"Free Hormone Estimate"** tests, abbreviated FT4E and FT3E.

(a) Nomenclature of FT4 and FT3 Methods

Considerable confusion encompasses the nomenclature of free thyroid hormone tests. Controversy continues regarding the technical validity of the measurements themselves and their clinical utility in conditions associated with binding protein abnormalities (52,137,140,142,143). Free hormone testing in clinical laboratories (FT4 and

FT3) are made either by index methods that require two separate tests, single test ligand assays or physical separation methods that isolate free from protein-bound hormone prior to direct measurement of hormone in the free fraction. Ligand assays are standardized either with solutions containing gravimetrically established concentrations of hormone, or use calibrators with values assigned by a physical separation method (i.e. equilibrium dialysis and/or ultrafiltration). Physical separation methods typically are manual, technically demanding and quite expensive for routine clinical use. Index and ligand assay FT4 and FT3 tests are more commonly used in the clinical laboratory setting, where they are typically performed on automated immunoassay analyzer platforms (17).

Unfortunately, a confusing plethora of terms have been used to distinguish the different free hormone methods and the literature is filled with inconsistencies in the nomenclature of these tests. Currently, there is no clear methodologic distinction between terms such as “T7”, “effective thyroxine ratio”, "one-step", "analog", “two step”, "backtitration", "sequential", “immunoextraction” or “immunosequestration”, “ligand assay” because manufacturers have modified the original techniques or adapted them for automation (140). Following the launch of the original one-step “analog” tests in the 1970s, the term “analog” became mired in confusion (140). This first generation of hormone-analog tests were shown to be severely binding-protein dependent and have since been replaced by a new generation of labeled-antibody “analog” tests which are more resistant to the presence of abnormal binding proteins (140,144). Unfortunately, manufacturers rarely disclose all assay constituents or the number of steps involved in an automated procedure so that it is not possible to use the method’s nomenclature (two-step, analog etc) to predict its diagnostic accuracy for assessing patients with binding protein abnormalities (144).

(b) Index Methods (FT4I and FT3I)

Indexes are free hormone estimate tests that require two separate measurements (138). One test is a total hormone measurement (TT4 or TT3) the other is an assessment of thyroid binding protein concentration using either an immunoassay for TBG, a T4 or T3 “uptake” test called Thyroid Hormone Binding Ratio (THBR). Alternatively, estimates of the free T4 fraction can be determined by isotopic dialysis or ultrafiltration. The quality and purity of the tracer employed critically impacts on the diagnostic accuracy of the indexes calculated using an isotopic signal (141,145,146).

(i) Indexes using TBG Measurement

Calculation of a FT4I using TBG only improves diagnostic accuracy compared with TT4 when the TT4 abnormality results from an abnormal concentration of TBG. In addition, the TT4/TBG index approach is not fully TBG independent, nor does it correct for non TBG-related binding protein abnormalities or for TBG molecules which have abnormal affinity (133,147-150). Thus, despite the theoretical advantages of using a direct TBG measurement, TT4/TBG indexes are rarely used because TBG binding capacity can be altered independent of changes in the concentration of TBG protein, especially in patients with NTI (103).

(ii) Indexes using Thyroid Hormone Binding Ratio (THBR) or “Uptake” Tests

“T3 uptake” tests have been used to estimate thyroid hormone binding proteins since the 1950s. Two different types of “uptakes” have been used. “Classical” uptake tests add a trace amount of radiolabeled T3 or T4 to the specimen and allow the labeled hormone to distribute across the thyroxine binding proteins in exactly the same way as the endogenous hormone (138,146). Since only a trace amount of labeled T3 or T4 is used, the original equilibrium is barely disturbed. The distribution of the tracer is dependent upon the saturation of the binding proteins. Addition of a secondary binder or adsorbent (anion exchange resin, talc, polyurethane sponge, or charcoal coated bead, etc.) results in a redistribution of the T3 or T4 tracer in a new equilibrium, which now includes the binder. The tracer counts sequestered by the adsorbent are dependent on the saturation of the binding proteins: the higher the saturation of the binding proteins, the greater the amount of tracer in the adsorbent. This results in an indirect measurement of FT4.
Under normal physiological conditions with normal amounts and affinities of binding proteins, high saturation of the thyroid binding proteins (especially TBG) results in a smaller amount of T3 tracer binding to the binding proteins and a larger amount being taken up by the adsorbent. Conversely, when TBG saturation is low, more T3 tracer becomes bound to the adsorbent. It has been recommended that a normal serum sample standard be

used to normalize the response of the assays and allow for the reporting of the result as a ratio to normal i.e. a "Thyroid Hormone Binding Ratio (THBR)" (146). Many manufacturers still use the "classical" approach to produce T3 uptake assays, in which the mean normal percent uptake can vary from 25% to 40% (bound counts/total counts). Traditionally, the free thyroxine index, sometimes called a "T7" is derived from the product of a T3-uptake test and a TT4 measurement, often expressed as a % uptake (adsorbent bound counts divided by total counts).

"Classic" T3-uptake or THBR tests are typically influenced by the endogenous T4 concentration of the specimen. This limitation can be circumvented by using a very large excess of a non-isotopically labeled T4 tracer with an affinity for thyroid binding proteins comparable to that of T4. This type of assay reflects the ***total*** binding capacity (TBC) of the sample, and is expressed as its ratio to the normal Binding Capacity assigned a ratio value of 1.0. With a TBC assay, the index is derived from TT4 divided by the Uptake ratio. The result is reported in ug/dL and the normal range is usually very similar to that of the TT4 assay. An FT4I calculated using the TBC approach usually correlates well with the T7 calculated using a percent T3- uptake assay, even though the two types of assays measure totally different aspects of thyroid hormone binding to the carrier proteins.

Current THBR tests usually produce normal FT4I and FT3I values when TBG abnormalities are mild (i.e. pregnancy). However, some of these tests will produce inappropriately abnormal index values when patients have grossly abnormal binding proteins (congenital TBG high or low, familial dysalbuminemic hyperthyroxinemia (FDH), thyroid hormone autoantibodies and NTI) and in the presence of some medications that influence thyroid hormone protein binding [Section 3B(e)].

B.2 Guidelines for Use of Thyroid Hormone Binding Ratio (THBR) or "Uptake" Tests

(e) "Uptake" tests should be called "Thyroid Hormone Binding Ratio" tests, abbreviated THBR and include an indication of which hormone is used, i.e. THBR (T4) or THBR (T3).
(f) A T4 signal is preferred over T3 for THBR measurements, to better reflect T4 binding protein abnormalities.
(g) THBR values should be reported as a ratio with normal serum, the latter having an assigned value of 1.00.
(h) THBR calculations should be based on the ratio between absorbent counts divided by the total minus absorbent counts, rather than the ratio between absorbent counts and total counts.
(i) The THBR result should be reported in addition to the total hormone and free hormone index value.
(j) THBR tests should not be used as an independent measurement of thyroid status, but <u>should</u> be interpreted in association with a TT4 and/or TT3 measurement and used to produce free hormone estimates (FT4 or FT3 indexes).

(iii) Measurements using Free Hormone Fraction

The first free hormone tests developed in the 1960s were indexes, calculated from the product of the free hormone fraction of a dialysate and a TT4 measurement (made by PBI and later RIA) (151,152). The free fraction index approach was later extended to measure the rate of transfer of isotopically-labeled hormone across a membrane separating two chambers containing the same undiluted specimen. The free hormone indexes calculated with isotopic free fractions are not completely independent of TBG concentration and furthermore are influenced by radiochemical purity, the buffer matrix and the dilution factor employed (153, 154).

(c) Ligand Assays for FT4 and FT3 Measurement

These methods either employ either a “two-step” or “one-step” approach. Specifically, two-step assays employ a physical separation of free from protein-bound hormone before free hormone is measured by a sensitive immunoassay, or alternatively, an antibody is used to immunoextract a proportion of ligand out of the specimen before quantitation. In contrast, one-step ligand assays attempt to quantify free hormone in the presence of binding proteins. Two-step methods are less prone to non-specific artifacts. One-step methods may become invalid when the sample and the standards differ in their binding of assay tracer (57,137,142).

(i) Ligand Assays employing Physical Separation

FT4 methods that physically isolate free from protein-bound hormone before employing a sensitive immunoassay to measure the free hormone concentration are standardized using solutions containing gravimetrically prepared standard preparations of T4. The physical isolation of free from protein-bound hormone is accomplished with either a semi-permeable membrane using a dialysis chamber, an ultrafiltration technique, or a Sephadex LH-20 resin adsorption column (153-157). An exceedingly sensitive T4 RIA method is needed to measure the picomolar concentrations of FT4 in dialysates or free fraction isolates, as compared with total hormone measurements that measure in the nanomolar range. Although there are no officially acknowledged "gold standard" free hormone methods, it is generally considered that methods that employ physical separation are least influenced by binding proteins, and by inference, provide free hormone values that best reflect the circulating free hormone status (97,158). However, dialysis methods employing dilution may underestimate FT4 when binding inhibitors are present in the specimen and adsorption of T4 to membrane materials may be an issue (97,158). In contrast, such methods may overestimate FT4 in sera from heparin-treated patients as a result of in-vitro generation of FFA (87,101,102,104,105,159-161). Physical separation methods are too labor intensive and expensive for routine use by clinical laboratories and are usually only available in reference laboratories. FT3 methods employing physical separation are only available in some specialized research laboratories (106).

(ii) Ligand Assays without Physical Separation

Most free hormone immunoassays in current use employ a specific, high affinity hormone antibody to sequester a small amount of total hormone from the specimen. The antibody occupancy, which is usually inversely proportional to the free hormone concentration, is quantified using a hormone labeled with radioactivity, fluorescence- or chemiluminescence-based reagent. The signal output is then converted to a free hormone concentration using calibrators with free hormone values assigned by a method employing physical separation. The actual proportion of total thyroid hormone sequestered (up to 5%) varies with the methodologic design, but greatly exceeds the actual free hormone concentration. The active sequestering of hormone by the anti-thyroid hormone antibody reagent in the assay results in a continuous re-equilibration of the bound/free equilibrium. The key to the validity of these methods is twofold. Firstly, it is necessary to use conditions that maintain the free to protein-bound hormone equilibrium, and minimize dilution effects that weaken the influence of any endogenous inhibitors present in the specimen. Secondly, it is important to use serum calibrators containing known free hormone concentrations that behave in the assay in an identical manner to the patient specimens. Three general approaches have been used to develop comparable FT4 and FT3 immunoassay methods: (i) two-step labeled-hormone; (ii) one-step labeled-analog; and (iii) labeled antibody.

B.3 Guidelines: For Manufacturers Developing Free Hormone Estimate Tests

(k) Methods that do not physically separate bound from free hormone should extract no more than 1-2% of the total hormone concentration hormone off the binding proteins, so that the thermodynamic equilibrium is maintained as much as possible.
(l) Minimize dilution effects that weaken the influence of any endogenous inhibitors present in the specimen.
(m) Use serum calibrators containing known free hormone concentrations that behave in the assay in an identical manner to the patient specimens.
(n) Perform the test procedure at 37°C.

• Two-Step, Labeled-Hormone/Back-Titration Methods

Two-step methods were first developed for research purposes in the late 1970s and were subsequently adapted to produce commercial FT4 and FT3 methods. During a first incubation step, these methods use a high affinity ($>1x10^{11}$ L/mol) anti-hormone antibody bound to a solid support (ultrafine Sephadex, tube or particles) to sequester a small proportion of total hormone from a diluted serum specimen. After a short incubation, unbound assay constituents are washed away before the second step in which sufficient labeled hormone is added to bind to all the unoccupied antibody-binding sites. After washing, the amount of labeled hormone bound to the solid-phase antibody is quantified relative to gravimetric standards or calibrators that have free hormone values assigned by a reference method. One-step labeled hormone-analog methods were introduced in the late 1970s. These new tests were less labor-intensive than two-step techniques. As a result, two-step methods waned in popularity despite comparative studies showing that they were less affected by albumin concentration and binding protein abnormalities that negatively impacted the diagnostic accuracy of the one-step analog tests (140,162-164).

• One-Step, Labeled Hormone-Analog Methods

The physicochemical validity of the one-step labeled hormone-analog tests hinged upon the development of a hormone analog with a molecular structure that was totally non-reactive with serum proteins, but could react with unoccupied hormone antibody sites. If these conditions are met, the hormone-analog, which is chemically coupled to a signal molecule such as an isotope or enzyme, can compete with free hormone for a limited number of antibody-binding sites in a classic competitive immunoassay format. Though conceptually attractive, this approach is technically difficult to achieve in practice, despite early claims of success. The hormone-analog methods were principally engineered to give normal FT4 values in high TBG states (i.e. pregnancy). However, they were found to have poor diagnostic accuracy in the presence of abnormal albumin concentrations, FDH, NTI, high FFA levels or with thyroid hormone autoantibodies. Considerable efforts were made during the 1980s to correct these problems by the addition of proprietary chemicals to block analog binding to albumin or by empirically adjusting calibrator values to correct for protein-dependent biases. However, after a decade of criticism, most hormone-analog methods have been abandoned because these problems could not be resolved (140).

• Labeled Antibody Methods

Labeled antibody methods also measure free hormone as a function of the fractional occupancy of hormone-antibody binding sites. This competitive approach uses specific immunoabsorbents to assess the unoccupied antibody binding sites in the reaction mixture. A related approach has been the use of solid-phase unlabeled hormone/protein complexes (sometimes confusingly called “analogs”) that do not react significantly with serum proteins, to quantify unoccupied binding sites on anti-hormone antibodies in the liquid-phase. The physiochemical basis of these complex-based labeled-antibody methods suggests that they may be as susceptible to the same errors as the older labeled-hormone analog methods. However, physicochemical

differences arising from the binding of analog to the solid support confers kinetic differences that results in decreased analog affinity for endogenous binding proteins and a more reliable free hormone measurement. The labeled antibody approach is currently the favored approach for automated free hormone testing.

(d) Performance of FT4 and FT3 Tests in Different Clinical Situations

The only reason to select a free thyroid hormone method (FT4 or FT3) in preference to a total thyroid hormone test (TT4 or TT3) is to improve diagnostic accuracy for detecting hypo- and hyperthyroidism for patients with thyroid hormone binding abnormalities that compromise the diagnostic utility of total hormone measurements (57). Unfortunately, the diagnostic accuracy of current free hormone methods cannot be predicted from either their methodologic classification (one-step, two-step, labeled antibody etc) or by vitro tests of their technical validity, such as performing a specimen dilution test. The index tests (FT4I and FT3I) as well as current ligand assay methods, are all protein dependent to some extent, and may give unreliable values when binding proteins are abnormal (52). Free hormone tests should be performed at 37°C since tests performed at ambient temperature falsely increase values when specimens have a very low TBG concentration (165,166).

B.4 Guidelines - Clinical Utility of Serum Free T3 Estimate Tests

Low serum T3 has little specificity or sensitivity for diagnosing hypothyroidism, since enhanced T4 to T3 conversion maintains serum T3 concentrations within normal limits until hypothyroidism becomes severe. Patients with NTI or caloric deprivation typically have low TT3 and FT3 values. Serum T3 measurements are usually interpreted together with FT4.

Serum T3 measurement is useful to diagnose complex or unusual presentations of hyperthyroidism and certain rare situations:

- A high serum T3 is often an early sign of recurrence of Graves' hyperthyroidism.
- A high or paradoxically normal serum T3 may indicate hyperthyroidism in an NTI patient with suppressed TSH (< 0.01 mU/L).
- A paradoxically normal serum T3 may indicate Amiodarone-induced hyperthyroidism.
- Patients with goiter living in areas of iodine deficiency often have FT3 measured in addition to TSH to detect T3 thyrotoxicosis caused by focal or multifocal autonomy.
- A high serum T3 is often seen with TSH-secreting pituitary tumors.
- A high serum T3 is often seen in thyroid hormone resistance syndromes.
- Serum T3 measurement is useful for monitoring compliance with L-T3 suppression therapy prior to 131I scan for DTC.
- Serum T3 measurement is useful for distinguishing mild (subclinical) hyperthyroidism (low TSH/ normal FT4) from T3-toxicosis.
- Serum T3 measurement is useful for investigating iodine deficiency (characterized by low T4/high T3).
- The T3/T4 ratio can be used to investigate the etiology of hyperthyroidism (Graves' disease versus non-Graves' disease). Excess thyroidal stimulation characteristic of Graves' disease gives rise to a high T3/T4 ratio (>20 ng/ug metric or >0.024 molar).
- Serum T3 measurement can be used to investigate whether a suppressed TSH in the presence of normal serum T4 is due to T3-toxicosis or mild (subclinical) hyperthyroidism.
- Serum T3 measurement can be useful during antithyroid drug therapy to identify persistent T3 excess, despite normal or low serum T4.
- Serum T3 measurement can be useful to diagnose amiodarone-induced thyrotoxicosis – a diagnosis which should not be based on T4 excess alone because of the frequency of euthyroid hyperthyroxinemia during amiodarone treatment.
- Serum T3 measurement can be used to monitor the acute response to treatment for Graves' thyrotoxicosis.
- Serum T3 measurement can be used to detect early recurrence of thyrotoxicosis after cessation of antithyroid drug therapy
- Serum T3 measurement can be used to establish the extent of T3 excess during high-dose replacement or

suppressive therapy with T4 or after an intentional T4 overdose.
- Serum T3 measurement can be useful when diagnosing thyroid hormone resistance which usually presents without clinical hyperthyroidism.

The impetus for developing free hormone tests has been the high frequency of binding-protein abnormalities that cause discordance between total and free thyroid hormone concentrations. Unfortunately, no current FT4 method is universally valid in all clinical conditions. When the concentration of TBG is abnormal, many FT4 methods give results that are more diagnostically useful than TT4 measurement.

Pre-analytical or analytical assay artifacts arise in many situations associated with binding protein abnormalities: when the binding of the assay tracer to albumin is abnormal; in the presence of medications that displace T4 from TBG; during critical phases of NTI; and in pregnancy (see Table 1). The frequency of these FT4 assay artifacts suggests that TSH or the TSH/FT4 relationship is a more reliable thyroid parameter to use than an estimate of FT4 alone. When it is suspected that a FT4 result is discrepant, FT4 should be checked using a different manufacturer's method (measured in a different laboratory). Additionally, or alternatively the FT4/ TT4 relationship can be checked for discordance since interference seldom affect both measurements to the same degree and in the same direction.

(i) Pregnancy

The increase in serum TBG and the low albumin concentrations associated with pregnancy results in wide method-dependent variations in FT4 measurements [see section 2A(c)] (40,54). Albumin-dependent methods can produce low FT4 values in up to 50 percent of patients and are unsuitable for use during pregnancy because of the marked negative bias attributable to the progressive decline in serum albumin concentration during pregnancy (54). Conversely, methods such as tracer dialysis that employ a high degree of specimen dilution tend to show a positive bias in relation to standards, secondary to the TBG-induced increase in the pool of protein-bound T4 (57).

B.5 Guidelines: Abnormal Thyroid Hormone Binding Proteins Effects on FT4 Tests

Binding protein abnormalities cause pre-analytical or analytical FT4 assay artifacts. Thyroid function should be assessed from the TSH-TT4 relationship when:

- The binding of assay tracer to albumin is abnormal (i.e. FDH).
- The patient is taking medications that displace T4 from TBG, i.e. Phenytoin, Carbamazepine or Furosemide (Frusemide).
- The patient has a critical or severe non-thyroidal Illness.

(ii) Premature Infants

Low thyroxine levels without elevated TSH is commonly encountered in premature infants of less than 28 weeks gestation (32,167). There is some clinical evidence to suggest that L-T4 treatment may improve neurological outcome (167). However, as described above, methodological differences in FT4 methods are likely to compromise the reliability of detecting hypothyroxinemia of prematurity.

(iii) Genetic Abnormalities in Binding Proteins

Heredity and acquired variations of albumin and TBG with altered affinity for either T4 or T3 can cause abnormal total hormone concentrations in euthyroid subjects with normal free hormone concentrations (133). The albumin variant responsible for familial dysalbuminemic hyperthyroxinemia (FDH) has a markedly increased affinity for T4 and numerous T4-analog tracers, resulting in spuriously high serum free T4 estimates with these tracers (137,168). In FDH, serum TT4 and FT4I values, as well as FT4 measured by analog-based methods, give supra-normal values, whereas serum TT3, FT3, TSH and FT4 measured by some two-step

methods or equilibrium dialysis are normal (168). Failure to recognize the presence of the FDH albumin variant that can occur with a prevalence of up to 1:1000 in some Latin-American populations can result in inappropriate thyroid ablation (169).

(iv) Thyroid Hormone Autoantibodies

Some patient sera contain autoantibodies to thyroid hormones that result in methodologic artifacts in both total as well as FT4 and FT3 measurements (135,137). Autoantibody interference is method-dependent. Tracer T4 or T3 bound to the endogenous antibody is falsely classified as bound in adsorption methods, or free in double antibody methods, leading to falsely low or falsely high serum TT4 or TT3 values, respectively (136,137). In particular, the T4 tracer analogs used in some FT4 tests may bind to these antibodies, leading to spuriously high serum FT4 results.

(v) Thyrotoxicosis and Hypothyroidism

The relationship between free and total T4 and T3 in thyrotoxicosis is non-linear. In severe thyrotoxicosis thyrotoxicosis when TT4 concentrations are two to three fold elevated, this non-linearity reflects both a decrease in TBG levels and an overwhelming of the TBG binding capacity despite increased binding to TTR and albumin (170). Similarly, FT3 concentrations may be underestimated as a result of high T4-TBG binding. The converse situation exists in severe hypothyroidism, in which there is reduced occupancy of all binding proteins (170). In this situation, an excess of unoccupied binding sites may blunt the FT4 response to treatment. This suggests that an initial L-T4 loading dose is the most rapid approach for restoring a therapeutic FT4 level in a hypothyroid patient.

(vi) Factors that Compete for Thyroid Hormone Binding

A number of medications and factors compete with the binding of T4 and T3 to TBG causing an acute increase in the availability of FT4 or FT3. This produces an inhibition in TSH secretion, an enhanced clearance of hormone and in some instances increased hormone action (171). Many competing agents of thyroid hormone binding are frequently prescribed therapeutic agents that differ in their affinity for TBG relative to T4 (100,172). Furosemide, for example binds to TBG but with an affinity that is about three-fold less than T4 whereas aspirin binds seven-fold less than T4 (173,174). The competition in vivo observed with such agents relates to their affinity for TBG rather than their therapeutic levels, free fraction or affinity for non-TBG proteins, especially albumin (173,175).

Current FT4 assays that employ a dilution factor may fail to detect an elevation in FT4 secondary to the presence of binding-protein competitors. For example, a specimen containing both T4 (free fraction 1:4000) and a competitive inhibitor (free fraction 1:100) subjected to stepwise dilution will sustain the FT4 concentration up to a 1:100 dilution, secondary to progressive dissociation of T4 from binding proteins. In contrast, the free drug concentration would decrease markedly only after a 1:10 dilution. Thus the hormone-displacing effect of drug competitors of T4 binding will be underestimated in FT4 assays employing the high specimen dilution. The use of symmetric equilibrium dialysis and ultrafiltration of undiluted serum can minimize this artifact (86,97,157,176).

(vii) Heparin Treatment Artifacts

It is well known that in the presence of a normal albumin concentration, non-esterified fatty acids (FFA) concentrations > 3mmol/L will increase FT4 by displacing the hormone from TBG (87,101,102,104,105,159-161,173). Serum from patients treated with heparin, including low-molecular weight heparin preparations, may exhibit spuriously high FT4 values secondary to in vitro heparin-induced lipase activity that increases FFA. This problem is seen even with heparin doses as low as 10 units and is exacerbated with storage of the specimen. Increased serum triglyceride levels, low serum albumin concentrations or prolonged assay incubation at 37°C can accentuate this problem.

(viii) Critical Nonthyroidal Illness

There is a large body of evidence collected for more than two decades, that report on the specificity of various FT4 methods in hospitalized patients with NTI [Section 2B(b)]. This literature can be confusing, and is complicated by the heterogeneity of the patient populations studied and the method-dependence of the results. Manufacturers have progressively modified their methods over time, in an attempt to improve their specificity in this setting and other situations when binding proteins are abnormal. However, the exact composition of current methods remains proprietary and it is difficult for manufacturers to obtain pedigreed specimens from such patients to rigorously test their methods. In one recent FT4 method comparison study, a marked method-dependent difference was seen on the seventh day following bone marrow transplantation in euthyroid subjects receiving multiple drug therapies that included heparin and glucocorticoids (105). In this studythe TT4 concentrations were normal in most of the subjects (95%) and the serum TSH was < 0.1 mIU/L in approximately half of the subjects, independent of the method. In contrast, both elevated and subnormal values were reported by different FT4 methods. It appears that the FT4 methods that employed specimen incubation at 37°C were prone to give supranormal FT4 estimates in 20 to 40% of patients, possibly due to the heparin effect noted above. In contrast, analog tracer methods that are subject to the influence of tracer binding to albumin, gave subnormal FT4 estimates in 20-30% of patients (105). Such FT4 measurement artifacts, giving rise to a discordance between FT4 and TSH results, increase the risk of an erroneous diagnosis of either thyrotoxicosis or secondary hypothyroidism and suggest that TT4 measurements may be more reliable in the setting of a critical illness.

(e) FT4 Method Validation

Unfortunately, most free hormone estimate methods receive inadequate evaluation prior to their introduction for clinical use. Manufacturers rarely extend the validation of their methods beyond the study of ambulatory hypo- and hyperthyroid patients, pregnant patients and a catchall category of “NTI/hospitalized patients”. However, there is currently no consensus as to the best criteria to use for evaluating these free T4 estimate methods. It is insufficient to merely demonstrate that a new method can distinguish between hypothyroid, normal and hyperthyroid values, and to show comparability with existing methods - any free hormone estimate method will satisfy these criteria without necessarily giving information about the true physiologic free hormone concentration.

New methods should either be tested with pedigreed clinical samples, especially those that may challenge the assay validity, or alternatively, by manipulating the constituents of a normal serum sample to test a particular criteria. Whichever approach is adopted, the key questions relate to the identity between samples and standards, because all assays are generally comparable. Other approaches include testing the quantitative recovery of added L-T4, or determining the effects of serum dilution (97). These approaches however, just test the “protein dependence” of the method, i.e. the degree to which free T4 is dependent on the dissociation of free from bound hormone (52). These approaches will predictably give an unfavorable assessment of methods that involve a high degree of sample dilution compared to those methods that minimize sample dilution. There is no evidence however, to document whether these approaches truly reflect diagnostic accuracy of the method when used to evaluate difficult clinical specimens. Ultimately, as with any diagnostic method, the specificity of a free T4 method will only become evident after testing a full spectrum of specimens from a non-diseased population. An unexpected interference may only be noted after methods have been in use for some time, as in the effects of rheumatoid factor that can produce spuriously high serum free T4 estimates (177). False fluorescence can be another cause of non-specific interference (178).

The preferred approach is to pay particular attention to specimens that are likely to cause non-specific interference in the assay result (102). Ideally, in the ambulatory patient setting these would include samples that have: a) TBG abnormalities (pregnancy, oral contraceptive therapy, and congenital TBG excess and deficiency); b) Familial Dysalbuminemic Hyperthyroxinemia (FDH); c) T4 and T3 autoantibodies and d), interfering substances such as rheumatoid factor and heterophilic antibodies (HAMA). In the hospital setting three classes of hospitalized patients should be tested: a) patients without thyroid dysfunction but with low or high TT4 due to NTI; b) patients with

documented hypothyroidism associated with severe NTI and, c) patients with documented hyperthyroidism associated with NTI. However, it is prohibitively difficult for manufacturers to obtain pedigreed specimens from such patients. Since no manufacturers have tested their methods adequately in the hospital setting, it is difficult for physicians to have confidence that abnormal FT4 results in such patients reflects true thyroid dysfunction rather than NTI. Thus in hospitalized patients with suspected thyroid dysfunction, a combination of serum TSH and TT4 measurements may provide more information than only a FT4 test. In ambulatory patients, serum FT4 measurements are often more diagnostically accurate than a TT4 measurement. When an abnormal FT4 result does not fit the clinical picture, or there is an unexplained discordance in the TSH to FT4 relationship, it may be necessary to order a TT4 test as confirmation. Alternatively, the laboratory could either send the specimen to a different laboratory that uses a different manufacturer's FT4 method, or to a reference laboratory that can perform a FT4 measurement using physical separation, such as equilibrium dialysis or ultrafiltration.

B.6 Guidelines for Manufacturer Assessment of FT4 Estimate Test Diagnostic Accuracy

- The diagnostic accuracy of the method should be tested using pedigreed specimens from ambulatory patients with the following binding protein disturbances:
  - TBG abnormalities (high estrogen & congenital TBG excess and deficiency)
  - Familial Dysalbuminemic Hyperthyroxinemia (FDH)
  - Increased Transthyretin (TTR) affinity
  - T4 and T3 Autoantibodies
  - Rheumatoid Factor
  - Heterophilic Antibodies (HAMA)
- Test the method for interference with normal serum specimens spiked with relevant concentrations of common inhibitors at concentrations that cause displacement of hormone from binding proteins in undiluted serum, effects which are lost after dilution i.e.:
  - Furosemide (Frusemide) 30 μM
  - Disalicylic acid 300 μM
  - Phenytoin 75 μM
  - Carbamazepine 8 μM
- List all known interferences with the magnitude and direction of resulting errors.
- Document in-vitro heparin effects on FFA generation during the assay incubation.

(f) Interferences with Thyroid Tests

Ideally, a thyroid hormone test should display zero interference by any compound, in any serum, at any concentration. Studies available from manufacturers vary widely in the number of compounds studied and in the concentrations tested. Usually the laboratory can only proactively detect interference from a "sanity check" of the relationship between the FT4 and TSH result. If only one test is measured, interference is usually first suspected by the physician observing an inconsistency between the reported value and the clinical status of the patient. Classic laboratory checks of verifying the specimen identity and performing after dilution, may not always detect interference. Interferences with either TT4 or FT4 measurements typically elicit inappropriately abnormal values in the face of a normal serum TSH level. (Table 1). Interferences with competitive or non-competitive immunoassays fall into four classes: (i) cross-reactivity problems, (ii) endogenous analyte antibodies, (iii) heterophilic antibodies and (iv) drug interactions (179).

(i) Cross-reactivity

Cross-reactivity problems result from the inability of the antibody reagent to discriminate selectively between analyte and a structurally related molecule (180). Thyroid hormone assays are less susceptible to this type of interference than TSH, because iodothyronine antibody reagents are selected for specificity by screening with

purified preparations. The availability of monoclonal and affinity-purified polyclonal antibodies has reduced the cross-reactivity of current T4 and T3 tests to less than 0.1% for all studied iodinated precursors and metabolites of L-T4.

(ii) Endogenous Autoantibodies

Endogenous Autoantibodies to both T4 and T3 have been frequently found in the serum of patients with autoimmune thyroid and non-thyroid disorders. Despite their high prevalence, interference caused by such autoantibodies is relatively rare. Such interferences are characterized by either falsely low or falsely high values, depending on the type and composition of the assay used (181).

(iii) Heterophilic Antibodies (HAMA)

HAMA fall into two classes (182). They are either relatively weak multispecific, polyreactive antibodies that are frequently IgM rheumatoid factor or may even be broadly reactive antibodies induced by infections or exposure to therapies containing monoclonal antibodies (177,183,184). These are sometimes called human anti-mouse antibodies (HAMA). Alternatively, they can be specific human anti-animal immunoglobulins (HAAA) that are produced against well-defined specific antigens following exposure to a therapeutic agent containing animal antigens (i.e. murine antibody) or by coincidental immunization through workplace exposure (i.e. animal handlers) (185). Either HAMA or HAAA affect IMA methodology more than competitive immunoassays by forming a bridge between the capture and signal antibodies, thereby creating a false signal, resulting in an inappropriately high value (186,187).

(iv) Drug Interferences

Drug Interferences can result from the in-vitro presence of therapeutic or diagnostic agents in the serum specimen in sufficient quanitities to interfere with the thyroid test (70,71). Thyroid test methods employing fluorescent signals may be sensitive to the presence of fluorophor-related therapeutic or diagnostic agents in the specimen (178). Alternatively, a therapeutic or diagnostic agent may have an in-vivo effect that disturbs the binding equilibrium of the thyroid hormones to the carrier proteins. In the case of I.V. heparin administration, the in vitro activation of lipoprotein lipases results in the generation of FFA in vitro which may falsely elevate FT4 values (87,101,102,104,105,159-161,173). Any competitor of hormone binding to serum proteins would be expected to initially elevate free hormone levels as a result of the reduced carrier protein availability. As the system re-establishes “normal” equilibrium, FT4 levels would normalize at the expense of a low TT4 concentration. The withdrawal of drug at this point would cause an initial fall in FT4 as more carrier protein becomes available, with re-normalization of FT4 as the equilibrium is re-established through an increased release of hormone from the thyroid gland. The time-scale and magnitude of competitor effects would differ with the half-life of the competitor agent.

(g) Serum FT4 and FT3 Normal Reference Intervals

Physical separation methods are used to assign values to the calibrators used for most FT4 estimate tests. There is closer agreement between the reference intervals of the various ligand assays used by clinical laboratories than there is between the various methods that employ physical separation. Reference intervals for FT4 immunoassay methods approximate 9-23 pmol/L (0.7 –1.8 ng/dL). In contrast, the upper FT4 limit for methods such as equilibrium dialysis that employ physical separation extends above 30 pmol/L (2.5 ng/dL). Reference intervals for FT3 immunoassay methods approximate 3.5-7.7 pmol/L (23 –50 ng/L). FT3 methods that employ physical separation are currently only available as research assays (106).

(h) Standardization or Calibration

There are no internationally developed standard materials or methods for free hormone measurement. Each method and manufacturer approaches the problem of standardization from its own unique perspective.

FT4 estimate methods that require two independent assays (tracer equilibrium dialysis and ultrafiltration as well as index methods) use a total hormone measurement and a measurement of the free fraction of the hormone. Total hormone assays are standardized with gravimetrically prepared calibrators from high purity hormone materials, which are commercially available. The free fraction is determined as radioactive counts in a dialysate or ultrafiltrate. Alternatively, in the case of the index methods, the saturation or binding capacity of the transport protein(s) is measured using a thyroid hormone binding ratio (THBR) test, sometimes referred to as an "uptake" test. THBR tests are standardized against sera with normal binding proteins and assigned a value of 1.00 [section 3B(b)ii].

The more complicated situation occurs with the ligand free hormone estimate assays. In general these tests are provided with standards that have known or assigned free hormone values determined by a reference method (usually equilibrium dialysis with RIA of the FT4 concentration in the dialysate). This is typically performed by the manufacturer for the purpose of establishing free hormone values for the human serum based calibrators containing the hormone and its binding protein(s) for inclusion in the kit. Alternatively, in the case of highly bound hormones, such as thyroxine, the Law of Mass Action can be used to calculate the free hormone concentration (188). The total hormone concentration, a measurement of the total binding capacity for the hormone in that serum sample, and the equilibrium constant provide the necessary information to calculate the free hormone concentration. This approach is valid for calibrators and controls manufactured in human serum that contains a normal TBG binding capacity. This allows the manufacturer to make calibrators and controls at fixed levels.

The use of calibrators, prepared as described above, also compensates for the over-extraction of hormone from their binding proteins. Specifically, in the case of thyroxine and triiodothyroninine, the antibody in the kit may bind the free hormone and extract some (~1-5%) of the bound hormone [Section 3B(c)]. If assayed directly, the concentration of free hormone would be elevated due to the over extraction. However, the use of calibrators of known free hormone levels and in human serum permits the assignment of the specific signal levels from the assay readout system (whether, isotopic, enzymatic, fluorescent, or chemiluminescent) to specific known concentrations of free hormone in a proportional relationship. This will only be valid provided that the percent of hormone extracted from the calibrator is identical to that from the patient specimen, even those containing binding protein abnormalities (i.e. congenital high and low TBG, FDH, NTI etc).

For a variety of reasons, there tends to be little numerical agreement in ranges for normal levels of free hormone across kits from multiple manufacturers and methodologies. This makes direct numerical comparison of data somewhat more difficult, since slopes will vary depending upon the two methods being compared. Therefore a better judgement can be made by looking at the agreement of the diagnosis of well-characterized patient samples measured across assay kits.

- Free Hormone Measurement – the Future

The era of immunoassay methods for the quantification of thyroid and steroid hormones in biological fluids began in the 1970's. That era is now drawing to a close. The era of advanced mass spectrometry for quantifying hormones in biological fluids is now emerging (131). There is no reason to doubt that mass spectrometry will provide better quantification because of greater analytical specificity and less analytical interference than immunoassays. However, for total hormone assays, the requirement for complete hormone release from protein-hormone complexes will remain. For free hormone assays, the requirement for a physiologic separation of free hormone and protein bound hormone, prior to quantification will also remain. In order to accomplish the later, new separation technology will be needed, before any methods can be regarded as the true gold standard. The implicit dilution of small molecules is a limitation of equilibrium dialysis that needs to be overcome. Ultrafiltration shows promise, but current methods are either too leaky or too impractical for this task. Mass spectrometry measurements of hormones that form complexes with serum proteins will only be as good as the specimen preparation steps associated with the quantification. However, the ideal free hormone reference method would be a technique that employs ultrafiltration at 37°C, to avoid dilution effects and the direct measurement of free hormone in the ultrafiltrate by mass spectrometry.

B.7 Guidelines for Laboratories Performing FT4 and FT3 testing

- Physicians should have ready access to information on the effects of drugs and the diagnostic accuracy of the test used for assessing the thyroid status of patients with various binding protein abnormalities and severe illnesses.
- When requested by the physician, the laboratory should be prepared to confirm a questionable result by performing a total hormone measurement or by re-measuring FT4 by a reference method that physically separates free from bound hormone, such as direct equilibrium dialysis or ultrafiltration.
- Questionable results on specimens should be checked for interference by re-measurement made with a different manufacturer's method. (Send out to a different laboratory if necessary.)