Taylor & Francis
Taylor & Francis Group

Critical Reviews in Clinical Laboratory Sciences

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ilab20

Free thyroxine measurement in clinical practice: how to optimize indications, analytical procedures, and interpretation criteria while waiting for global standardization

Federica D'Aurizio, Jürgen Kratzsch, Damien Gruson, Petra Petranović Ovčariček & Luca Giovanella

To cite this article: Federica D'Aurizio, Jürgen Kratzsch, Damien Gruson, Petra Petranović Ovčariček & Luca Giovanella (2022): Free thyroxine measurement in clinical practice: how to optimize indications, analytical procedures, and interpretation criteria while waiting for global standardization, Critical Reviews in Clinical Laboratory Sciences, DOI: 10.1080/10408363.2022.2121960

To link to this article: https://doi.org/10.1080/10408363.2022.2121960

Published online: 13 Oct 2022.

CRITICAL REVIEWS IN CLINICAL LABORATORY SCIENCES
https://doi.org/10.1080/10408363.2022.2121960

Taylor & Francis
Taylor & Francis Group

INVITED REVIEW

Free thyroxine measurement in clinical practice: how to optimize indications, analytical procedures, and interpretation criteria while waiting for global standardization

Federica D'Aurizio[a], Jürgen Kratzsch[b], Damien Gruson[c], Petra Petranović Ovčariček[d] and Luca Giovanella[e,f]

[a]Department of Laboratory Medicine, University Hospital of Udine, Udine, Italy; [b]Institute for Laboratory Medicine, Clinical Chemistry and Molecular Diagnostics, University Hospital, University of Leipzig, Leipzig, Germany; [c]Department of Clinical Biochemistry, Cliniques Universitaires St-Luc and Université Catholique de Louvain, Brussels, Belgium; [d]Department of Oncology and Nuclear Medicine, University Hospital Center Sestre milosrdnice, Zagreb, Croatia; [e]Clinic for Nuclear Medicine and Competence Center for Thyroid Diseases, Imaging Institute of Southern Switzerland, Ente Ospedaliero Cantonale, Bellinzona, Switzerland; [f]Clinic for Nuclear Medicine and Thyroid Center, University and University Hospital of Zurich, Zurich, Switzerland

ABSTRACT

Thyroid dysfunctions are among the most common endocrine disorders and accurate biochemical testing is needed to confirm or rule out a diagnosis. Notably, true hyperthyroidism and hypothyroidism in the setting of a normal thyroid-stimulating hormone level are highly unlikely, making the assessment of free thyroxine (FT4) inappropriate in most new cases. However, FT4 measurement is integral in both the diagnosis and management of relevant central dysfunctions (central hypothyroidism and central hyperthyroidism) as well as for monitoring therapy in hyperthyroid patients treated with anti-thyroid drugs or radioiodine. In such settings, accurate FT4 quantification is required. Global standardization will improve the comparability of the results across laboratories and allow the development of common clinical decision limits in evidence-based guidelines. The International Federation of Clinical Chemistry and Laboratory Medicine Committee for Standardization of Thyroid Function Tests has undertaken FT4 immunoassay method comparison and recalibration studies and developed a reference measurement procedure that is currently being validated. However, technical and implementation challenges, including the establishment of different clinical decision limits for distinct patient groups, still remain. Accordingly, different assays and reference values cannot be interchanged. Two-way communication between the laboratory and clinical specialists is pivotal to properly select a reliable FT4 assay, establish reference intervals, investigate discordant results, and monitor the analytical and clinical performance of the method over time.

Abbreviations: BK: binding capacity; BMI: body mass index; CLSI: Clinical & Laboratory Standards Institute; CV_A: inter-assay analytical variation; CV_G: between-subject biological variation; CV_I: within-subject biological variation; CV_P: preanalytical variation; DIT: diiodotyrosine; EuBIVAS: European Biological Variation Study; FT3: free triiodothyronine; FT4: free thyroxine; GPCR: G-protein coupled receptor; HAbs: heterophilic antibodies; ID: isotopic dilution; IFCC C-RIDL: International Federation of Clinical Chemistry and Laboratory Medicine Committee on Reference Intervals and Decision Limits; IFCC C-STFT: International Federation of Clinical Chemistry and Laboratory Medicine Committee for Standardization of Thyroid Function Tests; II: index of individuality; IS: internal standard; IVD: in vitro diagnostics; *K*: affinity; LC-MS/MS: liquid chromatography-tandem mass spectrometry; MIT: monoiodotyrosine; MRM: multiple reaction monitoring; NACB: National Academy of Clinical Biochemistry; NIS: sodium/iodide symporter; RCV: reference change value; RI: reference interval; RMP: reference measurement procedure; T3: triiodothyronine; T4: thyroxine; TBG: thyroxine-binding globulin; TFT: thyroid function tests; THAb: anti-thyroid hormone antibodies; THDP: thyroid hormone distributor protein; TPO: thyroid peroxidase; TSH: thyroid-stimulating hormone; TTR: transthyretin

ARTICLE HISTORY
Received 28 March 2022
Revised 13 July 2022
Accepted 3 September 2022

KEYWORDS
Thyroxine; free thyroxine; assay; standardization; thyroid dysfunctions

CONTACT Luca Giovanella luca.giovanella@eoc.ch Clinic for Nuclear Medicine and Competence Center for Thyroid Diseases, Imaging Institute of Southern Switzerland – Ente Ospedaliero Cantonale, Via A. Gallino 12, Bellinzona, 6500, Switzerland

Introduction

Thyroid dysfunctions are among the most common hormonal diseases. Laboratory evaluation is integral in the diagnosis and management of thyroid dysfunction, and thyroid function tests (TFT) are frequently ordered in both inpatient and outpatient settings. The thyroid-stimulating hormone (TSH) test is a reliable initial test with superior sensitivity and specificity compared with other thyroid hormone tests, in most cases [1]. Notably, serum TSH measurement is within the reference interval (RI) in most cases, especially in primary care, and further testing of thyroid hormones may not contribute to patient management [2,3]. However, measurement of free thyroxine (FT4) levels is pivotal in challenging conditions such as central hypothyroidism and hyperthyroidism, non-thyroid illness, and exogenous interferences. Moreover, FT4 assessment is needed to properly manage treated thyroid disorders. After TSH, FT4 is the most commonly ordered TFT, with approximately 18 million tests performed in 2008 in the United States compared with approximately 59 million TSH tests [1,4]. The accuracy of an FT4 test, however, is highly dependent upon the assay employed. The assays used in most clinical laboratories have some limitations and pitfalls. While the inter-assay precision of FT4 assays is generally good, the accuracy of that result may be poor. Indeed, in a survey of 13 FT4 methods, more than 50% of the results in four of the methods did not meet the allowable inaccuracy criteria [5]. To address this important issue, a working group for the international standardization of the FT4 assay was formed [6–8].

This review covers the clinical use of FT4, the characteristics and limitations of analytical methods for the measurement of FT4, and the role of standardization of FT4 assays in improving their clinical utility. Guidance for rational FT4 ordering and interpretation is also provided.

Thyroid physiology and pathophysiology

Under physiological conditions, thyroid follicular cells trap stable iodine by the sodium/iodide symporter (NIS), an intrinsic membrane protein that is part of the sodium/solute symporter family and the human solute carrier transporter family 5; it is located in the basolateral membrane of the follicular cells [9,10]. Iodine transport intracellularly is generated by the Na^+/K^+ ATPase pump, which provides the transmembrane Na^+ gradient. NIS transports one iodide ion together with two sodium ions and results in a significantly higher concentration of iodine in the follicular cells (up to 500 times) compared to the bloodstream. Subsequently, through different membrane channels such as pendrin, anoctamin, and the chloride channel, ClC5, located at the apical membrane, iodine passes from the cytoplasm of the follicular cell into the lumen [11,12]. Simultaneously, the glycoprotein thyroglobulin, in the process of exocytosis, passes the apical membrane and enters the follicular lumen. Thyroglobulin serves as the backbone for thyroid hormones and is the main component of the colloid in the thyroid follicular lumen; it is present in very high concentrations, up to 750 mg/mL [13]. The next step of iodine processing is oxidation, which is modulated by the enzyme, thyroid peroxidase (TPO). The process of oxidation is enabled by hydrogen peroxide, a substrate for TPO, that is synthesized at the apical border outside the follicular cell. Oxidation is followed by organification (i.e. incorporation of oxidized iodine by covalent links into tyrosyl residues of thyroglobulin), which enables the synthesis of diiodotyrosines (DITs) and monoiodotyrosines (MITs). One DIT and one MIT are then coupled in an oxidative process modulated by TPO to form triiodothyronine (T3), while two DITs form thyroxine (T4) (Figure 1) [14]. These hormones are phenolic rings coupled by an ether link and iodinated at three (3,5,3′-tri-iodo-L-thyronine, i.e. T3) or four (3,5,3′,5′-tetra-iodo-L-thyronine, i.e. T4) positions on the phenolic ring [13].

Thyroglobulin also serves as the storage receptacle of thyroid hormones in the lumen of follicular cells. If thyroid hormones are required, the thyroglobulin-thyroid hormone complex undergoes internalization to the cytoplasm by the process of endocytosis, as nonselective fluid uptake, or by receptor-related transport [15]. This is followed by enzymatic degradation and hydrolysis of the complex and transport *via* the basolateral membrane, mainly across the monocarboxylate transporter 8 (Figure 2).

Under physiological conditions, TSH regulates iodine uptake and the production of T3 and T4 with a positive linear TSH/radioactive iodine uptake relationship and an inverse log-linear TSH/FT4 correlation (Figure 3) [16,17]. This process is mediated by G-protein coupled receptors (GPCRs) at the basolateral follicular cell membrane. Upon binding of TSH, the activated TSH receptor dissolves heterotrimeric G protein. The released $G\alpha_s$ subunit activates adenylyl cyclase, and consequently, cyclic adenosine monophosphate accumulates in the cells. This pathway modulates the proliferation, differentiation, and function of thyroid cells, while regulators in the 5′ flanking region and transcription factors of the NIS genes modulate TSH-related transcription [18,19]. Through these pathways, TSH promotes transcription of

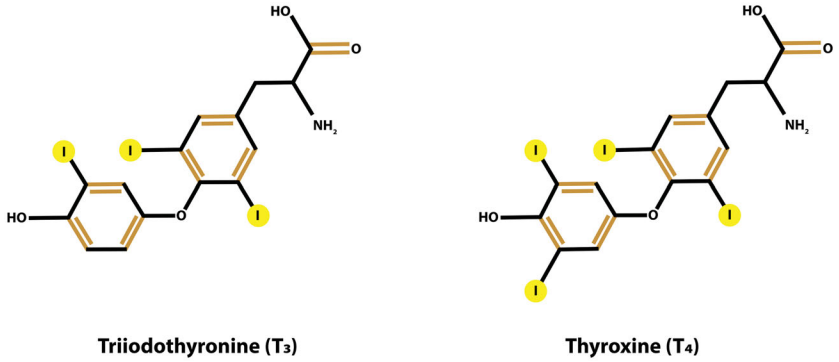

Figure 1. Chemical structure of thyroid hormones.

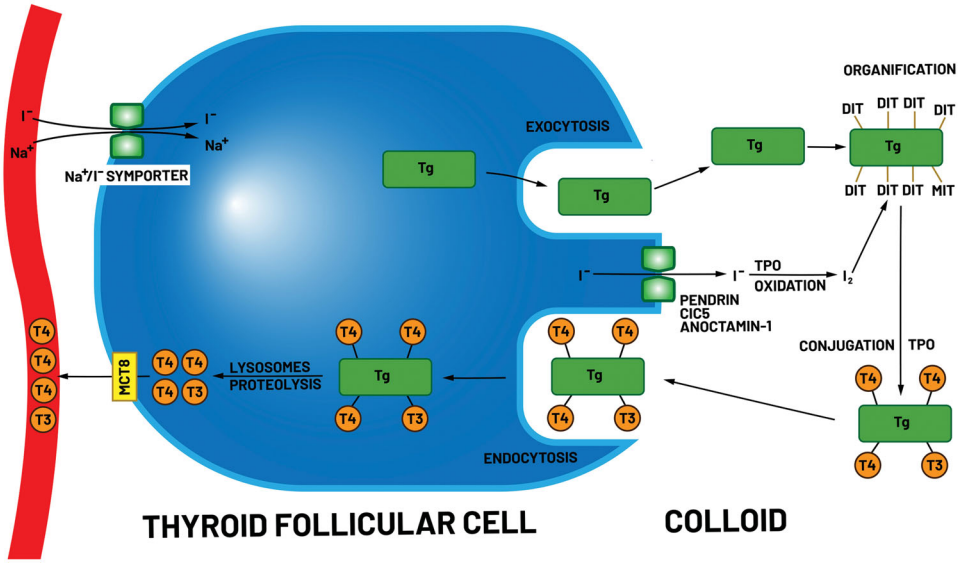

Figure 2. Biosynthesis of thyroid hormones. (T4, thyroxine, T3, triiodothyronine, MCT8, monocarboxylate transporter 8; Tg, thyroglobulin; TPO, thyroid peroxidase; DIT, diiodotyrosine; MIT, monoiodotyrosine).

the *NIS* gene, non-iodinated thyroglobulin exocytosis, transcription of genes for TPO, endocytosis of the iodinated form of thyroglobulin, the activity of the enzyme deiodinase, and T4 and T3 release to the bloodstream. Following activation by TSH, GPCRs are usually internalized. Subsequently, they undergo dephosphorylation, followed by recycling to the cell membrane. In some cases, GPCRs are degraded by lysosome enzymes [18].

Hormone production in the thyroid gland is focused mainly on the synthesis of T4 (85–90%) and, to a lesser extent, conversion of T4 to highly potent T3, which has four times higher activity than T4. In the thyroid, this process is regulated by the enzymes, 5′-iodothyronine deiodinases types 1 and 2. However, most of the T3 formation occurs not in the thyroid but in the peripheral tissues. Its production is driven in the liver by the activity of liver deiodinases. The biological half-life of T4 is approximately 1 week, while that of T3 is less than 24 h. Inactivation of T3 in peripheral tissues through the elimination of inner ring iodine is achieved primarily by type 3 iodothyronine deiodinase. The iodothyronine deiodinases that modulate the synthesis and deactivation of T3 are crucial for the maintenance of euthyroid status [20]. Both T4 and T3 are carried in the bloodstream by three major transporters: thyroxine-binding globulin (TBG), albumin, and transthyretin (TTR). TBG, a member of the serine protease inhibitor superfamily, has the highest affinity for thyroid hormones, while albumin has the highest concentration in the blood; however, albumin's affinity for T4 is 7000-fold lower compared with TBG. TBG binds up to 75% of serum T4 and 75% of serum T3, while albumin carries 5% of serum T4 and 20% of T3. The affinity of TTR for T4 is also 50-fold lower than that of TBG. TTR binds approximately 20% of serum T4 and less than 5% of serum T3.

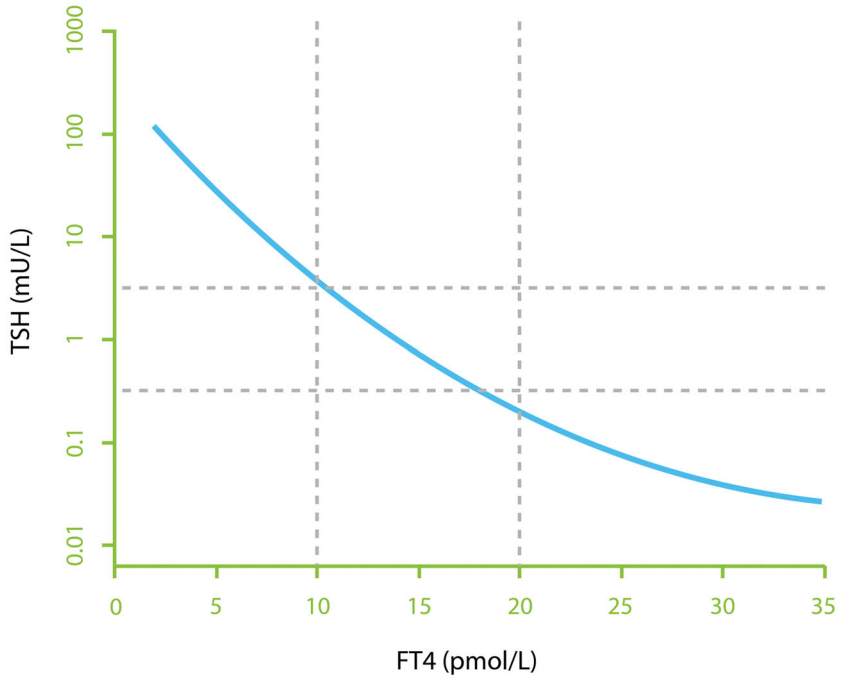

Figure 3. The log-linear inverse relationship between TSH and FT4. (TSH, thyroid stimulating hormone; FT4, free thyroxine). The blue line represents an approximate relationship between TSH and FT4. The dotted grey lines represent the normal values for TSH (horizontal) and FT4 (vertical).

However, T4 is much more firmly bound to carrier proteins compared with T3. Furthermore, less than 1% of these hormones circulate as free thyroid hormones (i.e. 0.03% serum T4 and 0.3% serum T3) [21].

Assessment of thyroid function: role and indications of for FT4 measurement

Symptoms of thyroid disease may be nonspecific, making laboratory diagnosis crucial. The assessment of thyroid dysfunction relies on the measurement of circulating concentrations of TSH, FT4, and, in some cases, free T3 (FT3). As mentioned above, TSH and FT4 have a complex, nonlinear relationship, such that small changes in FT4 result in relatively large changes in TSH [22]. Accordingly, the measurement of TSH is a sensitive screening test for thyroid dysfunction, and guidelines from the American Thyroid Association, the American Association of Clinical Endocrinologists, and the National Academy of Clinical Biochemistry (NACB) have all endorsed its measurement as the best first-line strategy for detecting thyroid dysfunction in most clinical settings [23–26]. However, despite the high negative predictive value of a normal serum TSH concentration in ruling out primary hypothyroidism or thyrotoxicosis, TSH alone is not appropriate for certain patient groups. In these instances, it is pertinent to test free thyroid hormones, primarily FT4, in addition to TSH [27,28].

Indications for FT4 measurement in clinical practice

Differentiation between subclinical and overt thyroid dysfunction

With the increased frequency of health screening and routine blood tests, many patients are now being diagnosed with subclinical thyroid dysfunction. It is important to note that even though the TSH test has been recommended as the first-line test for thyroid dysfunction, its sole utilization is insufficient in subclinical thyroid disease, and FT4 should also be tested. Briefly, normal FT4 in association with suppressed/reduced or increased TSH levels is observed in subclinical hyperthyroidism and subclinical hypothyroidism, respectively. In patients with subclinical hyperthyroidism, additional FT3 testing should be obtained to rule out T3 toxicosis, especially in iodine-deficient areas. The clinical impact of subclinical thyroid dysfunction depends on the degree of deviation of TSH. Many patients revert to a euthyroid status after 3–9 months [29,30]. Thus, in less severe cases, it is recommended that TFT may be repeated after a period of observation.

Secondary hyperthyroidism

Occasionally, patients may have an elevated FT4 level and an elevated or inappropriately normal TSH level at presentation. While interferences in laboratory assays may explain this constellation of TFT results, two other

rare differential diagnoses should be considered: secondary hyperthyroidism from a TSH-secreting pituitary adenoma (prevalence 0.85/1 million) and resistance to thyroid hormone (RTH; prevalence 1/40,000) [31], including RTH-β (due to thyroid hormone-β gene defects), RTH-α (due to thyroid hormone-α gene defects), and RTH of unknown etiology (likely due to a cofactor abnormality or an interfering substance). Patients with RTH-β have elevated FT4 and FT3 with normal or slightly elevated TSH. Patients with RTH-α have low FT4 and reverse T3, slightly elevated FT3, and normal or slightly elevated TSH. Patients can also undergo genetic testing to confirm the diagnosis. A high alpha subunit of TSH to whole TSH ratio is suggestive of a TSH-secreting tumor [32].

Secondary hypothyroidism

Patients with secondary hypothyroidism (prevalence 1/80,000–120,000) present with low FT4 and low or inappropriately normal TSH levels [33]. The history should be investigated for possible brain/pituitary surgery or radiotherapy, head injury/traumatic brain injury, severe postpartum hemorrhage, amenorrhea/infertility, and short stature. It is important to assess the other anterior pituitary hormones to rule out hypopituitarism before starting treatment for secondary hypothyroidism.

Non-thyroidal illness syndrome

The interpretation of TFT can be difficult in hospitalized patients and those recently discharged from the hospital [34]. Non-thyroidal illness syndrome is recognized as a nonspecific adaptive mechanism for illness and an indirect marker of disease severity in various conditions, including hospitalization in the critical care setting [35]. The underlying mechanisms include multiple and complex alterations (i.e. inhibition of iodide uptake and organification, suppression of thyroglobulin synthesis and reduction of thyroid hormone secretion, inhibition of the hypothalamus-pituitary-thyroid axis) mediated by cytokines such as interleukin-6 and tumor necrosis factor-alpha that have specific effects on the thyroid gland [36]. The hallmark of non-thyroidal illness syndrome is low FT3, with or without low FT4, in combination with normal or low TSH in clinically euthyroid patients. Furthermore, in the prolonged phase of critical illness, non-thyroidal illness syndrome is associated with adverse outcomes. In general, TFT are discouraged in these conditions whenever possible.

Treated hyperthyroidism and hypothyroidism

Among hyperthyroid and hypothyroid patients, the improvement in TSH levels tends to lag behind the improvement in free thyroid hormone levels during the initial phases of treatment [25,37]. The laboratory tests used to monitor hyperthyroidism depend on the type of treatment given. Serum TSH and FT4 (and FT3 in case of thyrotoxicosis) should be measured four weeks after the introduction of antithyroid drugs. Euthyroidism is confirmed by normalization of FT4 (and FT3) values. The test should be repeated, depending on the clinical situation. Little information can be gained by measuring TSH during this phase of treatment. During the maintenance phase, once euthyroidism has been obtained, FT4 (or FT3) measurements should be repeated (depending on the clinical situation) to tailor the dose of antithyroid drugs. The patient should be monitored clinically every year for two to three years after the end of treatment, with monitoring of TSH, FT4 (or FT3) levels carried out if there are any clinical abnormalities. Patients treated with radioactive iodine should be monitored every four to six weeks by measuring FT4 (or FT3) for the first three months of treatment. After this, monitoring will depend on the clinical situation. As the aim of treatment is to eradicate hyperthyroidism at the expense of hypothyroidism, in the short-medium term, TSH and FT4 should be monitored for three to six months following treatment. Annual monitoring of TSH levels is recommended to detect the recurrence of hyperthyroidism or long-term iatrogenic hypothyroidism [38,39].

In conclusion, while testing TSH alone is sufficient for general screening, both FT4 and TSH assays are needed to diagnose subclinical thyroid dysfunction and central hypothyroidism, investigate the effects of drugs, assess hospitalized patients, and accurately assess treatment effects (Table 1). For new cases and screening,

Table 1. Clinical indications for FT4 measurement.

Indications	
TSH suppressed	Differentiate subclinical from overt hyperthyroidism Assess the degree of hyperthyroidism
TSH increased	Differentiate subclinical from overt hypothyroidism Assess the degree of hypothyroidism
Hyperthyroidism treated with anti-thyroid drugs or radioiodine	Monitor the response (TSH unreliable in the initial months after therapy)
Pituitary diseases	Diagnosis of central thyroid dysfunctions (TSH unreliable)
Central hypothyroidism	Monitoring thyroxine therapy (TSH unreliable)

FT4: free thyroxine, TSH: thyroid-stimulating hormone.

TSH alone is tested first with reflex FT4 in case of TSH values <0.2 or >6.0 mIU/L (RI 0.4–4.0 mU/L), to minimize patient/clinician inconvenience and improve cost-effectiveness [40]. An important point about the measurement of FT4 (beyond whether the test is indicated) is the reliability of the result. If the measurement of FT4 is not indicated, for example, because the pretest probability is low, and the reliability of the results is limited, there is likely to be a substantial number of false-positive results.

Measurement of free thyroxine

The measurement of FT4 and, more generally, of free hormones, is based on the hypothesis of "the free hormone," according to which only the free, circulating fraction, or non-protein bound fraction, is available to reach target organs and carry out its biological activity. Therefore, over the years, clinicians have shown great interest in obtaining serum-free thyroid hormone concentrations for the diagnosis of thyroid disease [41]. In fact, blood FT4 concentrations may only be a "rough estimate" of the free hormone and be unable to predict the local free hormone complexity [42]. Once secreted by the thyroid gland, only a very low percentage of T4 (about 0.01%) circulates as the free form in the blood while most of it is bound to the transport proteins, TBG, albumin and TTR; also, a small percentage is bound to a variety of apolipoproteins [43]. FT4 concentrations are controlled by the equilibrium between the fraction of T4 bound to proteins and their free binding sites. This dynamic equilibrium is influenced by various mechanisms including the different binding affinities of the transport proteins, which can have the high binding capacity and low affinity and still carry many T4 molecules (i.e. albumin), or can have high affinity but low binding capacity (i.e. TTR and TBG).

The dynamic reserve is expressed by the mass action equation at equilibrium:

$$K = PBT \times T4/FT4 \times [Pfree]$$

This equation can be transformed into the following:

$$FT4 = PBT4/K \times [Pfree]$$

where *FT4* represents the free fraction of T4, *PBT* denotes the concentration of protein-bound T4, *K* is the affinity constant of the proteins toward T4, and *Pfree* is the concentration of the unbound free binding sites of the proteins. Using the mass action equation and knowing the concentration of FT4 and the binding proteins together with their affinity constants (*K*), it is possible to predict the concentration of FT4. On the basis of current knowledge, most experts believe that the different regulatory mechanisms result in an almost constant concentration of free hormone *in vivo* (and, at least initially, in blood samples collected for *in vitro* testing) and have applied this model to design assays for FT4 measurement. The development of FT4 methods, based on the free hormone hypothesis, began in the 1970s, thanks mainly to the original work of Ekins and colleagues [44]. Over the next three decades, a large number of methods have been employed, and, eventually, many were automated [45].

One concept of the free hormone hypothesis combines all of the approaches that have been developed over time to determine the concentration of FT4: an ideal valid assay should work without bias, despite any large variations (both absolute and relative) in the affinity and concentrations of serum T4-binding proteins [46]. To date, no routine tests satisfy this condition perfectly. The measurement of FT4 remains technically challenging because the vast majority of T4 is protein-bound, and any attempt to measure the free fraction inevitably disrupts the balance between free and bound hormones. Thus, the concept of the "window of validity" of an assay for the measurement of FT4 has become widespread. The ideal window of validity of a test is the physiological range of the serum concentration of binding proteins within which no serum used for that particular test is significantly affected by the T4 sampling process or by procedures that could otherwise affect the equilibrium [46].

The current methods used to measure FT4 are generally divided into direct methods that involve physical separation of the free fraction from the protein-bound hormone and indirect methods that are based on different immunoassay formats and selectively measure non-bound FT4 without disrupting the protein-bound T4 [47,48]. Indirect methods include formulas that calculate a free hormone "index" from a total hormone measurement corrected for the effects of binding proteins with a TBG measurement [49,50]. However, because indirect methods are strongly influenced by the concentration of transport proteins, in particular TBG, indirect methods with a free hormone index are obsolete and have been replaced by immunoassays [4].

Physical separation methods

Equilibrium dialysis and ultrafiltration are considered the gold standard methods for the measurement of free hormones such as FT4 [51]. These methods involve the separation of the free hormone from that bound to the proteins followed by measurement of the free hormone using a highly sensitive and specific assay [44,52].

Notably, the separation step requires careful evaluation because it is important that the equilibrium between the bound and free fraction of the analyte is unaltered [52].

In equilibrium dialysis, the serum sample is placed in a dialysis chamber where it is separated from an isotonic buffer solution by a semipermeable membrane capable of passing FT4 but not protein-bound T4. Over a period ranging from 12–24 h at 37 °C, the concentration of FT4 in the sample and the buffer reaches equilibrium. At this point, the FT4 concentration in the dialysate is measured with a sensitive immunoassay ("direct" equilibrium dialysis) or, more recently, by liquid chromatography-tandem mass spectrometry (LC-MS/MS). Measurement of FT4 in the buffer allows the calculation of the concentration of FT4 in the sample based on the total volume of the patient sample and buffer. "Direct" equilibrium dialysis methods using radioimmunoassay to measure FT4 in dialysate were described in the 1970s, with a commercial version being introduced in the 1990s. However, this product is no longer commercially available [53].

A lesser-used indirect FT4 measurement approach has also been described. In this case, a small amount of radiolabeled hormone is added to the sample before dialysis. When equilibrium is reached, the amount of radioactivity in the dialysate is measured. Along with the concentration of the total hormone that is measured separately, this fraction is used to estimate the free hormone concentration. A crucial condition of a reliable equilibrium dialysis method is that the FT4 concentration in the buffer compartment should be equal to the FT4 concentration of the undialyzed serum [43,45]. Factors that can compromise the achievement of this requirement include dilution, temperature, composition, and pH of the buffer, the time spent obtaining the equilibrium, and the effect of nonspecific binding (drugs, inhibitors, etc.). Among these factors, nonspecific binding plays an important role. In fact, nonspecific binding can disturb the balance between bound and free fractions and result in an underestimation of true FT4. Notably, no equilibrium dialysis method can be performed without dilution, and the buffer volume must be included in the dilution factor. Ideally, dialysis methods should be performed with the lowest possible dilution [51,54–56].

In ultrafiltration, a separation between the two fractions is achieved by applying centrifugal force to the sample. The serum sample is adjusted to pH 7.4, incubated for 20–30 min at 37 °C to reach binding equilibrium, and transferred to a centrifuge tube equipped with a semipermeable membrane and placed in a fixed position in the centrifuge. The application of centrifugal force pushes liquid and small molecules, including FT4, from the upper compartment that contains the sample to the lower compartment where the ultrafiltrate is collected. Thus, unlike in dialysis, the two phases are not in direct contact during ultrafiltration. The determination of FT4 in the ultrafiltrate is performed by an immunometric method or, more recently, by LC-MS/MS.

Although ultrafiltration techniques can be used for the determination of the free fractions of both thyroid and steroid hormones, there are currently no commercial kits available using this technique [57,58]. In general, the correlation between equilibrium dialysis and ultrafiltration methods has been reported to be very good. However, factors that can influence ultrafiltration methods only partially overlap with those described for equilibrium dialysis methods. Unlike in equilibrium dialysis, ultrafiltration is not susceptible to the potential problems (primarily dilution) associated with the use of buffer solution as no buffer is used. Similarly to equilibrium dialysis, the temperature at which the analysis is performed affects the free hormone concentration: an increase in the ultrafiltration temperature from 25 °C to 37 °C results in a 1.5-fold increase in the concentrations of FT4 [59]. Finally, the adsorption of T4 onto the membrane is another aspect to evaluate and avoid if possible. In fact, the loss of binding proteins during ultrafiltration can cause false increases in FT4. To reduce the possibility of potential protein loss, a balance must be found between the amount of centrifugal force and the type of membrane used [60]. Physicochemical methods also experience a higher rate of measurement errors than automated immunoassays, mainly due to defects in dialysis membranes or filtration devices [61].

In summary, as methods involving the physical separation of the free fraction of T4 (equilibrium dialysis or ultrafiltration) are complicated, time-consuming, and therefore not suitable for routine analysis, most laboratories use automated immunoassay approaches that have greater precision and higher throughput (though they do not always have great accuracy) [43].

Immunoassays

FT4 immunoassays provided by *in vitro* diagnostic (IVD) manufacturers have a competitive design because the small size of T4 prevents the use of a sandwich immunometric scheme. The latter requires analytes large enough (>1500–3000 Da) to allow simultaneous binding of two antibodies [43,61]. FT4 immunoassays can be based on the "two-step" or "one-step" principle, as described below (Table 2, Figure 4) [48,54].

Table 2. Main analytical characteristics of the most used FT4 immunoassays as quoted by the manufacturers.

Manufacturer/assay	Principle of immunoassay	LOD (pmol/L)	LOQ (pmol/L)	Assay range (pmol/L)	Imprecision (CV %) (intra-assay; inter-assay; total)	RIs[†] (pmol/L)
Abbott ARCHITECT Free T4	CMIA, two-step	3.60	5.15	5.15–64.35	2.3–5.3; ND; 3.6–7.8	9.01–19.05
Abbott Alinity i Free T4	CMIA, two-step	3.60	5.41	5.41–64.35	1.7–3.0; ND; 2.0–3.1	9.01–19.05
Beckman Coulter Access Free T4	CLIA, two-step	3.22	ND	3.22–77.20	1.8–4.4; 3.3–8.1; 4.3–9.2	7.86–14.41
Roche cobas Elecsys® FT4 IV	ECLIA, one-step with two sequential incubations	0.5	1.3	0.5–100.0	1.6–5.0; 1.9–6.3; ND	11.9–21.6
Siemens Healthineers Centaur FT4	CLIA, one-step with labeled analog	1.3	ND	1.3–155	2.2–3.3; 2.3–4.0; 3.4–4.6	11.5–22.7
Siemens Healthineers Atellica IM FT4	CLIA, one-step with labeled analog	1.3	ND	1.3–154.8	1.2–4.7; 2.2–6.8; ND	11.5–22.7

CMIA: chemiluminescent microparticle immunoassay; CLIA: chemiluminescent assay; CV: coefficient of variation; ECLIA: electrochemiluminescence assay; FT4: free thyroxine; LOD: limit of detection; LOQ: limit of quantification; ND: not disclosed; RI: reference interval.
[†]Reference intervals were calculated in a population of apparently healthy adult males and females. Information correct to May 2022.

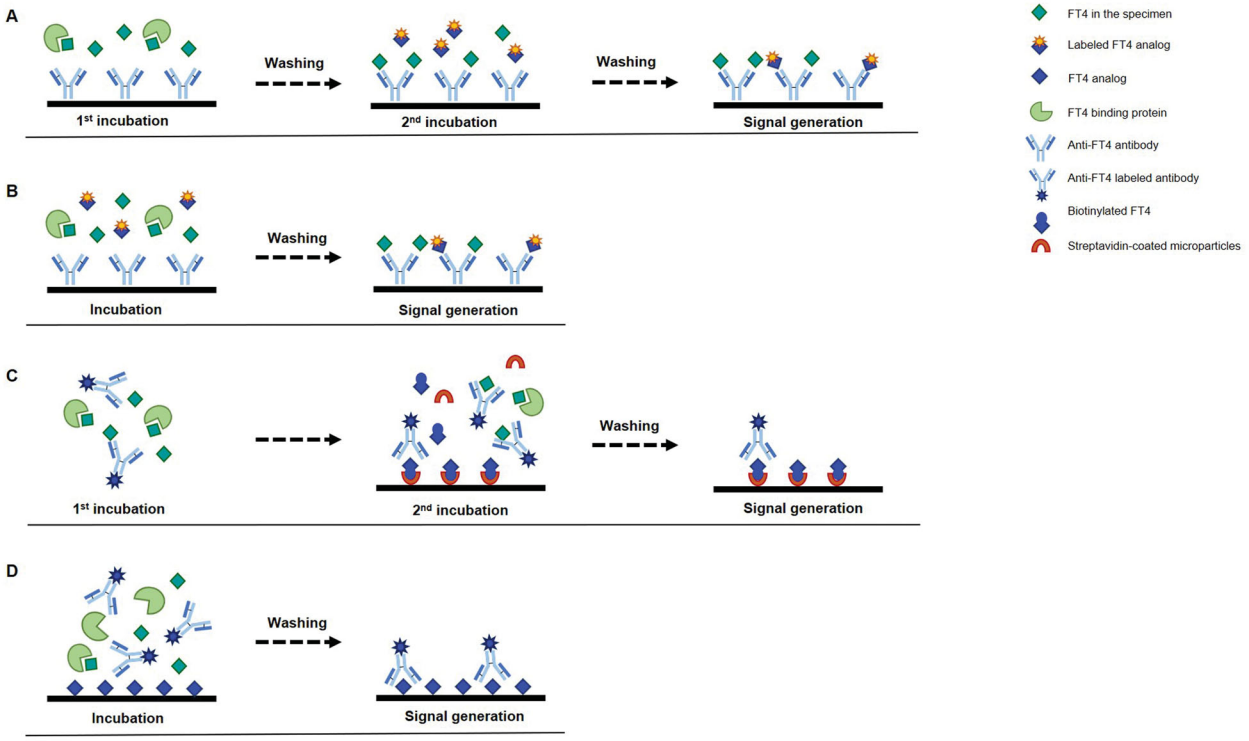

Figure 4. Schematic of the four main immunoassay formats used to measure FT4. (A) In the two-step labeled T4 analog method, FT4 in the specimen interacts with anti-hormone antibodies on a solid support. Antibody binding sites are occupied by endogenous FT4. The specimen is then removed and any empty antibody binding sites are allowed to interact with labeled FT4 ("back-titration"). (B) In the one-step assay, FT4 in the specimen competes with the labeled analog for binding to the antibodies fixed on a solid phase. The amount of labeled hormone left after the final washing of the solid support is inversely proportional to the original concentration of free hormone in the sample. (C) In the one-step analog two-step incubation format, anti-T4 antibodies are first incubated with FT4 in the specimen. In the second incubation step, the labeled T4 analog and streptavidin-coated microparticles bind to the unoccupied anti-T4 antibody binding sites, forming an antibody-hapten complex, bound to the solid phase via interaction between biotin and streptavidin. The unbound labeled T4 analog is washed away and the signal recorded. (D) In the labeled antibody method, labeled T4 antibodies bind to either FT4 in the specimen or solid phase-bound T4. After a short incubation period, a wash is carried out and the signal is measured (FT4, free thyroxine; T4, thyroxine).

Two-step assay (back titration)

The two-step assay was first described by Ekins and colleagues in the late 1970s (Figure 4A) [52]. In the first step, the serum is tested with anti-FT4 antibodies fixed on a solid support. The antibody sites are thus occupied in a manner directly proportional to the concentration of the free hormone present in the sample. After removing unbound serum constituents by washing, in the second step, a certain amount of labeled hormone (125-I T4) or, more recently, a macromolecular T4 conjugate, is added to occupy the remaining free sites. Once the excess labeled hormone has been removed, the binding sites of the antibody, which are inversely proportional to the concentration of free hormone originally present in the sample, can be quantified. This reaction scheme is termed "back-titration" [43].

The two-step approach has the advantage that the labeled T4 analog never comes into contact with the serum proteins (due to the washing phase immediately after capture) [4]. However, there is the potential negative impact of a concomitant loss of captured T4 antibodies before the tracer is added and the possible loss of bound T4 during competition with the tracer in the second incubation. Another key point to consider is dilution, which can alter the equilibrium between the endogenous FT4 and the protein-bound T4, weakening the robustness of the assay (the "phenomenon of sequestration") [43]. Finally, multiple steps can affect the analytical precision of the assay [4].

Some commercially available two-step assays/analyzers include the ARCHITECT Free T4 assay (Abbott, Chicago, IL, USA), Alinity i Free T4 Reagent Kit 200 Tests (Abbott, Chicago, IL, USA), UniCel DxI 800 Access Immunoassay System analyzer (Beckman Coulter, Brea, CA, USA) and DELFIA Immunoassays (PerkinElmer, Waltham, MA, USA) [62] (Table 2).

One-step assay (with analog)

The first component of the "one-step" approach employs a labeled hormone analog which competes with T4 in the sample for solid-phase anti-hormone antibody sites in a classic competitive immunoassay format (Figure 4B). After washing away the unbound constituents, the activity calculated from the binding between antibodies and the labeled analog is inversely proportional to the concentration of FT4 originally present in the serum. There is also a one-step analog two-incubation format (Figure 4C). In the first incubation, anti-T4 antibodies are mixed with the patient's serum and are immobilized in the solid phase. In the second incubation step, the labeled T4 analog and streptavidin-coated microparticles bind to the unoccupied anti-T4 antibody binding sites, forming an antibody hapten complex that is bound to the solid phase *via* interaction between biotin and streptavidin. The unbound labeled T4 analog is washed away, and the signal is recorded [63]. The key element of this one-step assay is the analog, which should be constructed to maintain its ability to interact with the analytical antibody but with structural chemical differences that do not theoretically allow binding to serum proteins.

The ideal situation described above has proven difficult to achieve. In fact, in the analogs, the alanine chain, which is responsible for binding with proteins, has been modified but it has not produced the desired reduction in binding affinity. It is estimated that the reduction in binding capacity is less than 50 times against TBG and thyroxine-binding prealbumin, and therefore is still far from a desirable level [64]. Although it is undeniable that this approach could offer substantial practical advantages, considerable criticism relating to the susceptibility of the assay to such factors as large changes in protein-bound total T4 concentration has been raised [65]. Thus, it was immediately observed that in a one-step assay with labeled analog, the results were satisfactory only when the physiological concentration of binding proteins was not altered, while significant limitations have been highlighted in pathophysiological conditions associated with abnormalities in binding proteins (i.e. pregnancy, familial dysalbuminemia, renal failure, etc.) or with thyroid-binding inhibitors such as those present in critical non-thyroidal illness syndrome [46,66,67].

Over the years, to at least partially overcome the problem, assay manufacturers have added albumin to the reagent to improve the behavior of the T4 analog in an effort (not always successful) to compensate for any alteration in the patient's serum protein concentration [46]. As described for the two-step assay (and the one-step assay with labeled analog), the phenomenon of sequestration has to be considered and avoided if possible. Thus, it is of paramount importance that the reaction between FT4 and the anti-T4 antibody occurs without removing a significant amount of the endogenous FT4 that is in equilibrium with the protein-bound hormone. Notably, to prevent disturbance of the equilibrium and thus maintain the accuracy of the assay, the reagent antibody should not remove more than 1% of the T4 in the sample. Scalar dilution has traditionally been used to measure the degree of susceptibility of an FT4 immunoassay to this effect, due to the amount and affinity of the reagent antibody. The assays are robust as long as the combined effect is minimized [45,68].

In the second and more recent version of the one-step assay, the tracer used is a labeled anti-hormone antibody (Figure 4D). FT4 fixed on a solid phase (e.g. magnetic beads) and FT4 in the serum sample compete for binding to the labeled antibody. After washing the unbound constituents away, the FT4 concentration is quantified as an inverse function of the fractional occupancy of hormone analog-labeled antibody binding sites in the reaction mixture [43,47,54]. The labeled antibody approach is used on several immunoassay platforms because it is easy to automate. In addition, it is considered less binding protein-dependent than the labeled analog approach because the solid phase hormone does not compete with endogenous free hormone for hormone-binding proteins. The key point is that the labeled antibody has a relatively low affinity for T4, which minimizes the disruption of the free/protein-bound equilibrium. Furthermore, the solid phase hormone analog usually has a macromolecular structure that is immobilized by the protein and therefore has different characteristics than the corresponding solution analogs of the other type of one-step assay: the steric hindrance of the labeled analog makes binding with the proteins present in the sample more difficult, although the macromolecular analog may still retain some capacity to bind to these proteins.

In addition to the problems highlighted above for both one-step approaches, further variables can make the results provided by these methods less reliable. In the most commonly marketed assays, various additives are used to influence the binding of the analogs with albumin, rendering the protein content less variable and ensuring the stability of ready-to-use reagents. These substances, although necessary from a chemical pretreatment point of view, may in many pathophysiological situations be able to disturb the complex relationship between the various components involved [64].

In summary, FT4 immunoassays have been criticized over the years for their poor diagnostic accuracy. Significant negative or positive biases have been described that exceed intra-individual biological variability, and assays demonstrate a weak FT4/TSH inverse log-linear relationship in several clinical conditions [16,69]. From the above, it is clear that although scientific research and technological innovation have contributed to a rapid evolution of methods for measuring FT4, there is still a need for further development that may lead to a new generation of easy-to-use and reliable analogs.

Some commercially available one-step assays/analyzers include the ADVIA Centaur XP immunoassay system (Siemens Healthineers, Tarrytown, NY, USA), Atellica Solution Immunoassay & Clinical Chemistry Analyzers (Siemens Healthineers, Tarrytown, NY, USA), AIA-900 Automated immunoassay analyzer (Tosoh Bioscience, Tokyo, Japan), cobas modular analyzer systems (Roche Diagnostics, Basel, Switzerland) [63], Vitros 5600 (Ortho Clinical Diagnostics, Rochester, NY, USA), MAGLUMI FT4 kit (Snibe Diagnostic, Snibe Co., Shenzen, China) [70], and Mindray FT4 (Mindray Bio-Medical Electronics, Shenzen, China) [62,71] (Table 2).

LC-MS/MS methods

To overcome the drawbacks of immunoassays, LC-MS/MS was proposed as an alternative approach to measuring total and free thyroid hormones (Figure 5) [42,61].

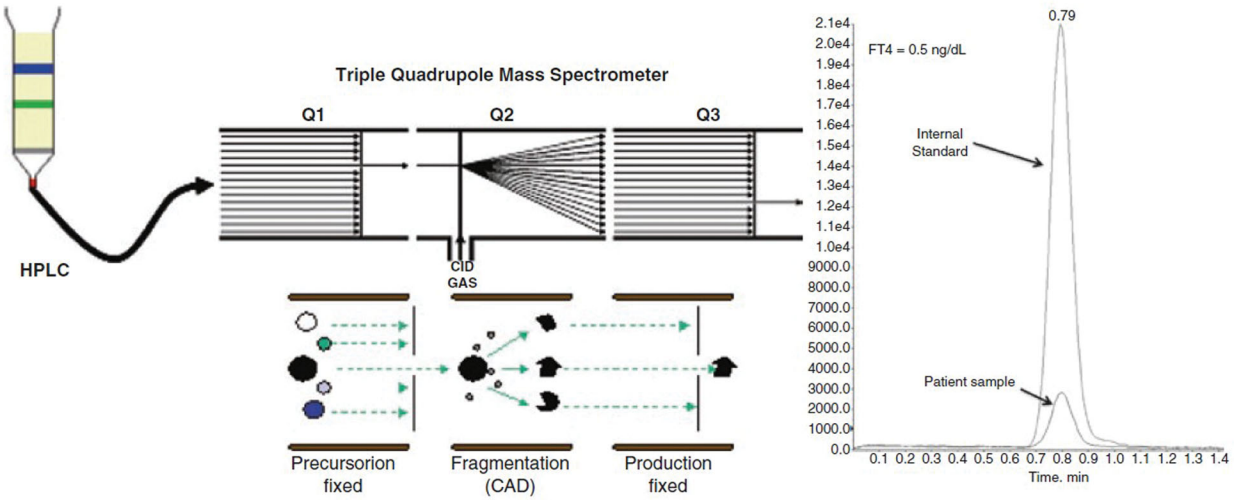

Figure 5. Schematic depiction of the workflow for current (high-pressure) liquid chromatography-tandem mass spectrometry (LC-MS/MS) of FT4.

LC-MS/MS methods identify the compound of interest based on both the retention time and the mass/charge ratio (m/z) of the precursor and of the ions produced by fragmentation, thus allowing greater analytical specificity and less analytical interference than immunoassays. In the early 2000s, Soldin et al. published the first method for determining free thyroid hormone concentrations in serum and plasma for diagnostic purposes [57]. The methods employed ultrafiltration of serum followed by LC-MS/MS. In detail, serum was filtered through an ultrafiltration device by centrifugation, labeled internal standard (IS, L-thyroxine-d2) was added to the ultrafiltrate, and a portion was injected onto a C-18 column. After washing, T4 and the IS were both eluted into a mass spectrometer system that utilized multiple reaction monitoring (MRM) quantification in the negative mode [57]. In the following years, Soldin's method was further improved by using equilibrium dialysis or ultrafiltration with centrifugation at 37 °C, different columns for elution, and more sensitive mass spectrometers [69,72,73].

The limited sensitivity of LC-MS/MS compared with immunoassays was initially a problem. Over the past decade, various optimizations have been completed, making the analysis of FT4 concentrations more practicable [42,74–76]. In addition, validated protocols that handle pre-analytical sample workup and subsequent LC-MS/MS analysis with the application of quality assessment parameters are being developed [77]. The improved sensitivity of modern MS instruments together with advanced phase chromatography allows the determination of total and free thyroid hormones. A recent publication that reported the first successful MS analysis of reverse T3 and 3,5-diiodothyronine along with T3 and T4 in human serum claimed an analytical sensitivity that reached the femtomolar level and showed a high degree of correlation with immunoassay techniques. However, this method still requires a serum volume of about 500 µL, which is not always available in clinical practice, especially in the intensive care or pediatric setting [78].

With LC-MS/MS measurements of FT4, a number of factors must be considered, in particular, matrix effect (ion suppression), derivatization or non-derivatization, and the types of ionization. Matrix effects occur because components of the matrix are co-extracted with the analyte of interest during the pre-analytical sample workup and co-elute during the chromatographic separation. This suggests possible interference in the efficiency of the analysis in MS with an increase ("ionic enhancement" by increase) or, more frequently for FT4, a decrease ("ion suppression") in signal strength. Ion suppression effects can cause interference in FT4 measurements. Ion suppression results from the presence of less volatile compounds that can modify the efficiency of droplet formation or droplet evaporation, which, in turn, affects the number of charged ions in the gas phase that ultimately reach the detector. Materials that can cause ion suppression include salts, ion coupling agents, endogenous compounds, drugs, metabolites, and proteins [79]. Depending on the strength of the matrix effect, even with the addition of IS to the sample, the accuracy of FT4 quantification can be impaired, especially at low concentrations. Matrix effects can be minimized by improving sample purity prior to analysis, using gradients that separate interferents from the analyte of interest, and evaluating IS peak (signal) heights, which should remain more or less constant among the samples [41]. A serious shortcoming of many publications reporting the quantification of thyroid hormones in biological samples, including human serum, is the lack of adequate elimination of matrix effects [42].

Pre-analytical derivatization has been proposed by some experts to improve the sensitivity of FT4 detection. However, other experts have highlighted the disadvantages, which include a reduction in precision due to additional derivatization steps (extraction, derivatization at one extreme of pH). Furthermore, derivatization methods are time-consuming, rendering this technique less convenient for adoption in routine clinical laboratories. To date, no procedure for routine analysis of FT4 in human serum or other biological specimens has been established. However, in recent years, the detection limits of modern mass spectrometers have improved significantly [57,69,80].

Numerous studies have compared immunoassays and LC-MS/MS for FT4 measurement in serum samples. In general, the correlation is acceptable in the euthyroid population but decreases significantly in clinical conditions that are characterized by changes in thyroid hormone transport proteins. Recently, in a group of patients admitted to an intensive care unit, Welsh and colleagues showed that LC-MS/MS and immunoassay results differed markedly [81]. This difference could be due to serum hypoproteinemia and/or the administration of drugs known to cause artifacts in thyroid hormone immunoassays, particularly in the free fractions [82–84]. Moreover, Araque et al. highlighted the importance of LC-MS/MS in the measurement of FT4, particularly in elderly patients with subclinical or overt hypothyroidism: the inverse log-linear correlation between TSH and FT4 was significantly better when FT4 was determined with LC-MS/MS compared with

immunoassays [85]. In considering the analytical performance of LC-MS/MS versus immunoassays for the testing of free thyroid hormones, the American Thyroid Association guidelines for thyroid management disorders during pregnancy recognize LC-MS/MS as the gold standard, although Anckaert et al. demonstrated that at least two current FT4 immunoassays were able to give FT4 patterns during pregnancy that were similar to those obtained by equilibrium dialysis isotopic dilution (ID) mass spectrometry [86,87]. Although LC-MS/MS methods have been shown to provide better quantification due to higher analytical specificity and less analytical interference than immunoassays, some experts have criticized the suggestion that LC-MS/MS be considered the gold standard for FT4. They assume that FT4 measurements by LC-MS/MS are valid only if the sample is equally valid. Thus, methods that use equilibrium dialysis or ultrafiltration to separate the free fraction from the protein-bound hormone are not without technical problems as described above; dilution, absorption, membrane defects, temperature, pH, and sample-related effects can cause non-uniformity between methods in LC-MS/MS, with some factors being more problematic than others in certain clinical scenarios [41]. Therefore, new separation technologies are needed before any method can be considered a gold standard.

Although LC-MS/MS methods are thought to be very reliable, they can produce erroneous data. A fundamental initial requirement is the unique identification of the target analyte in complex biological matrices. For each analyte of interest, at least two ionic transitions from the precursor ion to the ion fragment should be monitored, one for quantification, and the other for confirmation. The relative responses of these transitions (MRM) remain specific and consistent for each molecule. Possible MRM signals arising from any fragments of interfering molecules can be safely excluded from quantification. Unfortunately, many published and currently used methods in the thyroid hormone field do not apply those principles and use only a single transition to measure one analyte of interest or use two MRMs but without confirmation of the ion ratio [69]. Currently, most publications related to LC-MS/MS methods describe parameters supported by datasets and only a few provide complete documentation of the preanalytical and analytical quality assessment parameters used for hormone determination. However, the broad spectrum of new applications of LC-MS/MS-based analysis of thyroid hormones and metabolites in various biological samples highlights the future potential of this new technology. There are a number of important limitations to incorporating thyroid hormone measurement by LC-MS/MS in a routine setting. First, the technique requires the use of specialized and expensive equipment. Secondly, trained and dedicated personnel are needed to develop and implement the LC-MS/MS method, interpret the resulting complex data sets, and continuously monitor quality parameters. The cost of thyroid hormone measurement by LC-MS/MS is higher than that by immunoassays performed on automated platforms. Finally, LC-MS/MS has a longer response time and cannot provide fast results in emergency settings [75].

Standardization: facts, problems, and perspective

Despite some improvements over time, FT4 measurements of the same specimens by different immunoassay platforms continue to differ [5,8,67,88–92]. Assay variations affect FT4 RIs with potential clinical implications in reporting and interpreting results. In fact, patients are often referred to more than one laboratory, as laboratories use different methods for their FT4 assay, and physicians working in separate facilities may discuss the results of clinical cases without realizing that different methods have been used. In addition, over time, new measurement methods are introduced, and laboratories may change the method they use, for example, when technical supplies are replaced [48]. The common assumption that laboratories can eliminate method differences by adjusting their reference intervals according to the method used has not been confirmed. RIs vary substantially between laboratories, and there is no clear correlation between the measurement procedure used and the suggested RIs. The most likely reasons are that laboratories use different sources of information and different study designs when establishing and validating their RIs. The diversity of the RIs can influence the interpretation of the results and have implications for the clinical treatment of patients [93,94].

The International Federation of Clinical Chemistry and Laboratory Medicine Committee for Standardization of Thyroid Function Tests (IFCC C-STFT) developed and established a reference measurement system for FT4 standardization [4,8,95,96] and is now working with national partners on implementing it [97]. First, in collaboration with the IVD industry, the IFCC C-STFT initiated a quality and comparability assessment of commercial FT4 immunoassays, which were all found to be of good quality but demonstrated differences in test results [98]. In particular, using samples from patients with thyroid disease and exploring the potential impact of

standardization across the clinically relevant FT4 concentration range, all 13 FT4 assays tested showed bias in the medium to high concentration range, with the maximum discrepancy between assays being approximately 30% in the medium to high concentration range and approximately 90% in the lower concentration range [95]. The primary goal of FT4 immunoassay standardization is to establish metrological traceability and to ensure that analytical results are comparable between assays and laboratories over time. The standardization process requires a reference measurement system to use as part of the a calibration hierarchy; then, calibrated assays used in the different laboratories provide measurements attributable to the upper part of the hierarchy within the declared uncertainty limits [99]. A reference measurement system typically consists of (i) reference measurement procedures (RMPs) that are traceable to the highest reference available as outlined in ISO 17511 [100], and (ii) reference materials intended for use as calibrators and accuracy checks, with target values assigned by the RMP. The IFCC C-STFT has developed an international RMP using equilibrium dialysis for measuring serum FT4 concentrations, which is calibrated with a certified pure T4 standard and therefore traceable to the International System (SI, Système International) (pmol/L at a physiological pH of 7.40 and temperature of 37 °C) [7,101,102]. The RMP for FT4 proposed by C-STFT is indicated as conventional, with convention referring to the equilibrium dialysis part of the RMP, which must stringently observe a predefined procedure. In fact, although equilibrium dialysis is performed under ideal physiological conditions, it is not possible to prove unambiguously that the concentration of T4 in the dialysate is equal to the true concentration of FT4 in the original sample [101]. The implementation of the reference system for FT4 represents the most critical step in standardizing FT4 measurements [97]. It must be maintained and used in a transparent, consistent, and sustainable way to obtain reliable and comparable results. Well-documented procedures that describe how the reference system is used to calibrate and evaluate measurements at the lower levels of the traceability chain, such as those used by test manufacturers. must be in place. In addition, data such as RIs are needed to facilitate the use of standardized measurement procedures. Such requirements are addressed more efficiently with formal international standardization programs through national and regional organizations.

The reference system for FT4 is maintained through a network of laboratories using conventional FT4 RMPs. The initial members of the network were located at the University of Ghent, Belgium; the US Centers for Disease Control and Prevention; Radboud University Medical Center, Nijmegen, The Netherlands; and the Japanese Reference Material Institute for Clinical Chemistry Standards. The network of laboratories, which is now expanding, ensures the consistency of reference measurements over time and creates the laboratory capacity necessary to meet the demands for reference measurements. Further activities of other members of the reference laboratory network to establish and verify the correct implementation of metrological traceability, as required in the recently published ISO 17511 standard, are foreseen [100]. Changes in FT4 measurements after standardization will be relevant to many FT4 immunoassays. In this regard, a recent study by C-STFT has shown that although recalibration of FT4 tests is feasible and produces more comparable results between tests, it will result in an average variation of the measurement results of approximately 30–50% [95,96]. Undoubtedly, the use of immunoassays for FT4 will necessitate the establishment of new RIs [96]. The IFCC Committee on Reference Intervals and Decision Limits (C-RIDL) has recognized various obstacles to harmonizing RIs and identified different strategies to overcome these obstacles [103]. Current projects include a study to compare alternative approaches for the determination of FT4 RIs; a website to provide RIs obtained from a global study conducted by C-RIDL for the practice of evidence-based laboratory medicine; and a publication on the distinction between RIs and clinical decision limits. Reference value studies should be performed in different populations, including healthy adults, pregnant women, individuals taking levothyroxine, children, and infants [48]. Age-appropriate RIs will also be needed, particularly in children and adolescents. Manufacturers will be required to compare RIs in studies using current methods and the C-STFT reference method. In addition, the recalibration equation for IVD manufacturers will need to be certified and monitored over time as a guarantee of stability. Regulatory requirements associated with test recalibration will also need to be considered before standardization can be implemented. C-STFT has reached out to major regulatory agencies to determine what they will require from IVD manufacturers who recalibrate their tests. It should be noted that the re-approval of FT4 immunoassays by several health authorities (e.g. the US Food and Drug Administration) is a very broad process that requires the use of substantial resources from manufacturers [97]. After standardization, the clinical decision limits used in practice guidelines will require reevaluation and new guidelines for FT4 testing will be needed. Experts recognize the risk that changes to numerical measurement values and RIs after standardization could lead to

misinterpretation of laboratory data. Therefore, a major challenge in the standardization process will be to educate and prepare laboratories, clinicians, and patients for the new FT4 methods and changes in FT4 results [48]. To better address these concerns, the C-STFT announced these changes to various stakeholders and gathered information from IVD manufacturers, clinical laboratories, physicians, and individuals from IFCC member organizations on potential concerns, communication needs, and the benefits of FT4 standardization. The results of these comprehensive structured and informal assessments have been summarized on the C-STFT website [104].

The general recommendation of experts is that education on FT4 standardization should take place at three levels and should be supervised by an international working group led by the IFCC: (i) guidelines and expert recommendations published in journals; (ii) congress communications; and (iii) laboratory communications in the local community [48]. National and international scientific societies for laboratory medicine and clinical endocrinology will play a key role in enabling discussion, sharing promotion of standardization, explaining the changes and illustrating the differences in FT4 numerical values before and after standardization [97]. It is important that these educational programs also include information on what standardization will not achieve (e.g. they will not address the effects of binding proteins or interferences due to factors present in individual patient samples, including biotin). IVD manufacturers will also play a crucial role in implementing change in a coordinated way. Ideally, they should have responsibility for providing literature to laboratories to explain the standardization process in terms of why it is needed, when the changes will take effect, and how the transition will be managed. It would also be desirable for IVD companies to facilitate user group meetings to help laboratory professionals understand the changes and to enable them to communicate information to healthcare professionals [97].

In conclusion, the work of the C-STFT project has contributed greatly to progress in the field of recalibration of FT4 testing. However, much effort is still needed before global standardization can be achieved. Recalibration as a result of the project has had a considerable impact on the tested assays; inter-assay differences have been eliminated, and the remaining data scatter is almost completely due to within-assay error.

The current RMP for FT4 testing is technically challenging. Work is in progress among reference laboratories to address technical problems and optimize procedures [105]. Implementation challenges include defining clinical decision boundaries in different patient populations and educating all stakeholders. Experience gained from previous standardization programs has given valuable insight into the potential problems that can arise and planning strategies to overcome them. Without the strong involvement of clinical societies and the adoption of clinical guidelines and standards in the endocrinology community, education and acceptance of standardized FT4 values will not work [106].

Result interpretation

The classification of test results as "abnormal" or different from previous results is based on RIs and the reference change value (RCV).

FT4 reference intervals

The adequate interpretation of RIs for both TSH and FT4 is essential for the correct diagnosis and management of thyroid disease. In fact, the use of inappropriate RIs can influence clinical decision-making and have potentially harmful effects on the quality of patient care, including misdiagnosis, delayed diagnosis, and inappropriate treatment [107–109]. By definition, RIs describe the distribution of observed results in an apparently healthy reference population and thus are used to classify the test result in a patient as "normal" or "abnormal," and direct clinicians to interpret laboratory data in the context of the patient's overall clinical assessment [110,111]. Stringency is important in defining a reference population, as stated by the NACB, but this can be difficult and expensive to achieve [26]. Individuals free from thyroid disease history (personal, family, drugs) with normal physical exam (no goiter, normal thyroid ultrasound) and absence of thyroid antibodies are required. Many studies have been conducted over the years to determine FT4 RIs [48,109], a good proportion of which used the direct method according to the Clinical Laboratory Standards Institute (CLSI) EP28-A3c guideline, *Defining, establishing, and verifying reference intervals in the clinical laboratory* [108,109]. Individuals from a healthy population (the reference population) are selected on the basis of defined criteria. Samples are then collected from these individuals and analyzed for the selected measurand. A designation of good health for a referral candidate can be done using medical history, physical examination, and/or laboratory tests. Parametric or non-parametric statistics are then applied to the results obtained to determine the reference values [108]. The two main goals of any study aiming to set an FT4 RI are to compare the results obtained in the investigators' local population with those provided by the manufacturer

and to determine RIs in particular groups (pediatric, geriatric, pregnant individuals, and patients taking levothyroxine). Obtaining RIs using a direct method can be particularly challenging due to the technical difficulty of including sufficient numbers of healthy reference individuals. For example, in the infant population, samples with a small volume of blood together with ethical considerations that limit re-sampling contribute to the difficulty of establishing RIs that accurately describe the continuing physiological changes occurring in childhood [112]. In recent years, thanks to the growing interest in data mining, indirect approaches with methods such as Hoffman, Bhattacharya, and Arzideh have emerged to overcome these difficulties [107].

The main rationale for using indirect methods is that clinical pathology databases contain thousands or millions of results from hundreds or thousands of patients, including results from patients without the disease. Thus, using clinical exclusion and statistical approaches, reference distributions can be identified [113–115]. Some advantages of the indirect approach compared with the direct approach are the lower cost, faster results, less effort relating to ethical problems such as obtaining informed consent, and uniformity of preanalytical and analytical techniques used. The main disadvantages of the indirect approach lie in the need for advanced statistical methods to manage a large amount of data and in the clinical hypothesis that most of the individuals included are healthy [116].

All in all, as recommended by the CLSI EP28-A3C guideline, the responsibility of the evaluation of FT4 RIs to ensure their validity for local populations is delegated to individual laboratories. However, it is clear that most laboratories do not have the resources to make such an assessment and rely on the RIs provided by manufacturers [108,117].

FT4 reference intervals in adults

Examination of literature data from the last 16 years (January 2005 to December 2021) reveals that the type of immunometric method used to measure FT4 greatly influences the RIs, highlighting the urgent need for standardization of these assays (Tables 3–6; [117–145]) [121,146]. However, variation has been seen even when using the same method [145]. A recent study by Lee et al. showed a difference of 14.3% and −6.1% between the manufacturer and the direct determination for the lower and upper reference limit of FT4, respectively [48,147]. Such discrepancies may result from the study design, including the criteria used to enroll subjects: age, sex, iodine status, body mass index (BMI), medication, ethnicity, and geographic location [148,149].

The criteria for selecting an appropriate target population were addressed by the NACB and CLSI [108,126]. However, many published studies did not strictly adhere to the criteria indicated by international guidelines and few of them have evaluated thyroid normality with ultrasound. In addition, even the studies that used the aforementioned criteria reported FT4 RIs that were not always comparable between studies (Tables 2–5). To this end, Barth et al. showed that different intervals were found on two separate occasions even when identical criteria were used [145].

Ethnicity has been evaluated as a factor that can influence differences in RIs even when using the same method. A recent systematic review analyzed several studies that described differences in the concentration of analytes including TSH and FT4 between ethnic groups in both adult and pediatric populations [149]. However, the main results of studies looking at ethnicity differences were often not consistent with one another; thus the clinical significance of ethnic differences in relation to FT4 concentration remains unclear for now [150].

In addition to the analytical method and population selection criteria, the development of FT4 RIs also depends on sample collection (e.g. timing of sampling, handling of samples, storage) and the data mining technology and statistical tests used [117]. Little analysis of the statistical aspect has been done thus far. In one of the few studies on this subject, Strich et al. examined the effect of different statistical tools on RIs obtained with the data mining approach in a local population. The authors used different normalization techniques for the removal of outliers and obtained clearly different RIs depending on the statistical approach used [151].

It should be noted that regardless of the analytical method used, a statistically significant difference in FT4 RIs between sexes was demonstrated in just under half of the studies we analyzed, with higher FT4 values in males than in females (Tables 3–6) [109,120,125,130,133,135,138,139,142,143]. Differences in FT4 RIs between age groups have also been reported (Tables 3–6), paying special attention to the elderly (Table 7; [122,124,137,152]) [127]. In particular, current research has focused on the elderly. In this population, the correct determination of thyroid hormones is of paramount importance as elderly individuals often do not present noticeable signs and symptoms that allow the timely diagnosis of thyroid-related diseases [153–155]. Moreover, some studies have shown that TSH and FT4 levels in euthyroid elderly men are were associated with survival outcomes [156–158]. Cross-sectional studies have reported mixed results regardless of the analytical platforms employed; thus, some authors concluded that the

Table 3. Comparison of RIs of FT4 in adult populations on Abbott ARCHITECT platform.

Reference	Country	Number of patients and enrollment criteria	Partitioning criteria	Method for RI calculation	FT4 eRI (pmol/L)	FT4 mRI (pmol/L)
Quinn et al. [118]	China	414 (217 M; 197 F) Age range: 20–63 years Inclusion: TPOAb and TgAb negativity	Sex	Direct	Overall: 10.7–17.1	9.0–19.1 (De Grande et al. [121])
Schalin-Jäntti et al. [119]	Finland	511 (269 M; 242 F) Age range: 18–91 years Inclusion: TPOAb negativity	Age	Direct	Overall: 10.9–16.8	
Milinkovic et al. [120]	Serbia	22,860 (11,440 M; 11,420 F) Age ≥18 years Inclusion: TPOAb and TgAb negativity	Age; sex	Indirect	Overall: 10.5–18.9 M: 10.8–18.3* F: 11.5–15.4*	
Ehrenkranz et al. [122]	USA	69,223 (22,765 M; 46,458 F) Age range: 1–104 years Inclusion: 0.4 < TSH < 10 mIU/L Exclusion: personal or family history of thyroid disease; medications altering thyroid function; obstetric events within 9 months	Age; sex; daily and annual variations	Indirect	Overall: 9.8–19.1†	
Hickman et al. [123]	Australia	1177 (630 M; 547 F) Age range not specified Inclusion: TPOAb and TgAb negativity Exclusion: medications altering thyroid function	Age	Direct	Overall: 10.8–16.8	
Barth et al. [117]	UK	214 (83 M; 131 F) Age range: 18–65 years Inclusion: TPOAb and TgAb negativity Exclusion: pregnant or lactating women; blood donors; individuals on medication or with long-term conditions such as diabetes	Sex	Direct	Overall: 10.6–15.5	
Barhanovic et al. [124]	Serbia	790 F Age range: 20–69 years Inclusion: 0.1 < TSH < 10 mIU/L Exclusion: personal or family history of thyroid disease; medications altering thyroid function; obstetric events within 9 months; breastfeeding	Age	Direct	20–29 y: 10.7–17.5 30–39 y: 10.6–16.1 40–49 y: 10.4–18.1 50–59 y: 10.7–18.1** 60–69 y: 11.0–19.1**	

eRI: experimental reference interval, F: female, FT4: free thyroxine, M: male, mRI: manufacturer reference interval, RI: reference interval, TgAb: anti-thyroglobulin antibodies, TPOAb: anti-thyroid peroxidase antibodies, TSH: thyroid-stimulating hormone, UK: United Kingdom, USA: United States of America, y: years.
†Mean of age and sex results.
*Statistically significant difference between M and F and between 31–40 and 41–50 years old.
**Statistically significant difference (higher values) in the two oldest cohorts compared with the younger groups.

concentration of FT4 was normal or slightly reduced in elderly individuals [124,136,137,140,152] while others suggested that FT4 was slightly or not obviously increased [119,122,128,134,138,155,159]. Therefore, to reach reliable and consistent conclusions, further studies that systematically evaluate all possible factors influencing FT4 values in old age are needed.

FT4 reference intervals in pregnancy

Maternal thyroid dysfunction during pregnancy increases the risk of pregnancy-related complications including miscarriage, placental abruption, intrauterine growth restriction, premature birth, low birth weight, and impaired fetal neurodevelopment [86,160–163]. Diagnosis of thyroid dysfunction during pregnancy can be difficult as changes in thyroid physiology and altered serum matrices influence TSH and FT4 immunoassays. The state of the maternal thyroid must be interpreted in the context of the dynamic changes in thyroid production and metabolism related to pregnancy, the thyrotropic effect of β-human chorionic gonadotropin, the estrogen-induced increase in TBG, the augmented clearance of iodine by the kidneys, and the acceleration of T4 clearance by placental deiodinase

Table 4. Comparison of RIs of FT4 in adult populations on Beckman Coulter platforms.

Reference	Country	Platform	Number of patients and enrollment criteria	Partitioning criteria	Method for RI calculation	FT4 eRI (pmol/L)	FT4 mRI (pmol/L)
Plouvier et al. [125]	France	Access 2	262 (72 M; 190 F) Age range: 18–65 years Selection according to NACB guidelines (Baloch et al. [126])	Sex	Direct	Overall: 8.5–14.9 M: 8.6–14.9* F: 8.4–14.6*	7.9–14.4 (De Grande et al. [121])
Wang et al. [127]	China	DxI 800	4095 (2439 M; 1656 F) Median age: 43 years Exclusion: overt diseases; pregnancy	Age; sex	Indirect	M: 8.9–14.7** F: 8.6–14.3**	
Barth et al. [117]	UK	DxI 800	253 (105 M; 148 F) Age range: 18–65 years Inclusion: TPOAb and TgAb negativity Exclusion: pregnant or lactating women; blood donors; individuals on medication or with long-term conditions such as diabetes	Sex	Direct	Overall: 7.9–14.0	
Li et al. [128]	China	DxI 800	313 (173 M; 140 F) Age range: 20–97 years Inclusion: TPOAb and TgAb negativity Exclusion: pregnant or lactating women; personal or family history of thyroid disease, abnormal liver function, tumor, or other endocrine and metabolic disease	Sex	Direct	Overall: 8.6–14.3 M: 9.0–14.5 F: 8.5–14.3	
Wang et al. [129]	China	DxI 800	1828 Age range: 21–73 years Inclusion: TPOAb and TgAb negativity; sonographically assessed normal thyroid gland	Sex	Direct	Overall: 8.9–15 M: 9.1–15.2 F: 8.8–14.5	

eRI: experimental reference interval, F: female, FT4: free thyroxine, M: male, mRI: manufacturer reference interval, NABC: National Academy of Clinical Biochemistry, RI: reference interval, TgAb: anti-thyroglobulin antibodies, TPOAb: anti-thyroid peroxidase antibodies, UK: United Kingdom.
*Statistically significant difference between M and F.
**Statistically significant difference between M and F; statistically significant difference between younger and older than 50 years old.

Table 5. Comparison of reference intervals of FT4 in adult populations on Roche Diagnostics platforms.

Reference	Country	Platform	Number of patients and enrollment criteria	Partitioning criteria	Method for RI calculation	FT4 eRI (pmol/L)	FT4 mRI (pmol/L)
Kratzsch et al. [130]	Germany	Elecsys	453 (279 M; 174 F) Age range: 18–68 years Inclusion: healthy blood donors, with normal thyroid gland assessed sonographically Exclusion: personal history of a thyroid disease	Age; sex	Direct	Overall: 12.8–20.4 M: 13.0–21.4* F: 12.2–20.0*	
Friis-Hansen et al. [131]	North Europe (DK, FI, NO, SE)	Modular E170	430 (226 M; 204 F) Age ≥18 years Inclusion: TPOAb and TgAb negativity Exclusion: chronic disease; blood donation; pregnancy or breastfeeding; intake of prescription drugs (except contraception pills)	Age; sex; country	Direct	Overall: 14.5–22.9	
Takeda et al. [132]	Japan	Elecsys	878 (451 M; 427 F) Age range: 20–86 years Selection according to NACB guidelines (Baloch et al. [126])	Age; sex	Direct	Overall: 14.2–21.0	
Chan et al. [133]	China	Modular E170	157 (66 M; 91 F) Age ≥18 years Inclusion: TPOAb and TgAb negativity Exclusion: consumption of tobacco and alcohol; acute, chronic, or congenital illnesses and/or infections; history of malignancies; use of herbs or vitamin supplements; limb amputation; overnight shift duties; heavy manual workers; pregnancy, recent miscarriage, or menstrual disturbance in female participants; personal or family history of thyroid disease; medications	Age; sex	Direct	Overall: 13.2–20.7 M: 13.5–21.3* F: 12.6–19.7*	12.0–22.0 (De Grande et al. [121])
Yoshihara et al. [134]	Japan	cobas	1388 (198 M; 1190 F) Age range: ≥20 years Inclusion: normal thyroid gland assessed sonographically; negativity of anti-thyroid antibodies Exclusion: personal or familiar history of thyroid disorders; pregnant women; admission to hospital within the previous month; medication that might affect thyroid function; history of extrathyroidal malignancy; collagen disease; diabetes mellitus; neurologic disease; hepatic disorder; cardiovascular disease; chronic kidney disease	Age; sex	Direct	Overall: 11.7–20.0 M: 12.9–21.1 F: 11.6–19.7	
Amouzegar et al. [135]	Iran	cobas	2199 (953 M; 1246 F) Age ≥20 years Selection according to NACB guidelines (Baloch et al. [126])	Age; sex	Direct	Overall: 11.7–19.9 M: 12.2–20.6* F: 11.5–19.3*	

(*continued*)

Table 5. Continued.

Reference	Country	Platform	Number of patients and enrollment criteria	Partitioning criteria	Method for RI calculation	FT4 eRI (pmol/L)	FT4 mRI (pmol/L)
Marwaha et al. [136]	India	Elecsys	1916 (916 M; 1000 F) Age ≥18 years Inclusion: normal thyroid gland assessed sonographically; TPOAb negativity Exclusion: history of thyroid disease; family history of thyroid disease; use of medications known to interfere with thyroid function; systemic illness; goiter	Age; sex	Direct	Overall: 10.1–24.8**	
Fontes et al. [137]	Brazil	Modular E170	1200 (600 M; 600 F) Age ≥18 years Exclusion: past or present thyroid disease; thyroid surgery; family history of thyroid disease; TSH < 0.1 mIU/L or > 10.0 mIU/L; goiter; smoking; medicines known as possible analytical or physiological interference on measurement of TSH or FT4 in the past 3 months	Age; sex	Direct	M: 1.9–21.9 F: 1.9–24.5	
Sriphrapradang et al. [138]	Thailand	cobas e 411	1947 (1047 M; 900 F) Age ≥14 years Exclusion: pregnant women; history of thyroid disorder; moderate to severe illness; medications affecting thyroxine-binding proteins or thyroid function	Age; sex; residential region	Direct	Overall: 12.4–22.9	12.0–22.0 (De Grande et al. [121])
Mirjanic-Azaric et al. [139]	Serbia	cobas	227 (73 M; 154 F) Age range not specified Selection according to NACB guidelines (Baloch et al. [126])	Age; sex	Direct	Overall: 12.3–20.0*	
Park et al. [140]	Korea	cobas	5987 (3104 M; 2883 F) Age ≥10 years Inclusion: TPOAb negativity Exclusion: thyroid disease; family history of thyroid dysfunction; current pregnancy	Age; sex	Indirect	Overall: 11.8–20.6***	
Barth et al. [117]	UK	cobas	233 (92 M; 141 F) Age range: 18–65 years Inclusion: TPOAb and TgAb negativity Exclusion: pregnant or lactating women; blood donors; individuals on medication or with long-term conditions such as diabetes	Sex	Direct	Overall: 12.5–19.6	

DK: Denmark, eRI: experimental reference interval, F: female, FI: Finland, FT4: free thyroxine, M: male, mRI: manufacturer reference interval, NO: Norway, RI: reference interval, SE: Sweden, TgAb: anti-thyroglobulin antibodies, TPOAb: anti-thyroid peroxidase antibodies, TSH: thyroid-stimulating hormone.
*Statistically significant difference between M and F.
**Statistically significant difference between 18–30 years and >60 years in females and 18–30 years and >71 years in males.
***Statistically significantly decrease in FT4 with age.

Table 6. Comparison of reference intervals of FT4 in adult populations on the Siemens ADVIA Centaur platform.

Reference	Country	Number of patients and enrollment criteria	Partitioning criteria	Method for RI calculation	FT4 eRI (pmol/L)	FT4 mRI (pmol/L)
Reix et al. [141]	France	379 (192 M; 187 F) Selection criteria according to NACB guidelines (Baloch et al. [126])	Age; sex	Direct	Overall: 10.5–18.9	11.5–22.7 (De Grande et al. [121])
Wang et al. [142]	China	211 (94 M; 117 F) Selection criteria according to NACB guidelines (Baloch et al. [126])	Age; sex	Direct	Overall: 11.7–18.9*	
Cai et al. [143]	China	717 (330 M; 387 F) Selection criteria according to NACB guidelines (Baloch et al. [126])	Age; sex	Direct	Overall: 11.0–20.4 M: 11.4–21.1* F: 10.9–19.7*	
Hoermann et al. [144]	Germany	271 (68 M; 203 F) Exclusion: hypothalamic/pituitary diseases; pregnancy; use of anti-thyroid drugs or thyroid replacement regimes	Age; sex	Direct	Overall: 11.1–17.3	
Barth et al. [145]	UK	721 (A: 219; B: 222; C: 280) A: Redundant serum samples from primary care B: Repetition of series A after an interval of 12 months C: Healthy individuals	Not applied	Indirect (A and B); direct (C)	Overall: 10.0–20.0 A: 10–18.1 B: 11.1–20.6 C: 11.8–19.2	
Barth et al. [117]	UK	253 (99 M; 154 F) Age range: 18–65 years Inclusion: TPOAb and TgAb negativity Exclusion: pregnant or lactating women; blood donors; individuals on medication or with long-term conditions such as diabetes	Sex	Direct	Overall: 11.8–19.0	
Zou et al. [109]	China	20,303 (10,170 M; 10,133 F) Median age: 37 years Healthy individuals extracted from the hospital and laboratory information system	Age; sex	Indirect	Overall: 12.2–20.1 M: 12.8–20.6* F: 11.9–18.9*	

eRI: experimental reference interval, F: female, FT4: free thyroxine, M: male, mRI: manufacturer reference intervals, NACB: National Academy of Clinical Biochemistry, RI: reference interval, TgAb: anti-thyroglobulin antibodies, TPOAb: anti-thyroid peroxidase antibodies, UK: United Kingdom.
*Statistically significant difference between M and F.

Table 7. Comparison of RIs of FT4 in elderly populations using different analytical platforms.

References	Analytical platform	Age range (years)	Number of patients	FT4 cRI (pmol/L)	FT4 eRI (pmol/L)
Ehrenkranz et al. [122]	Abbott ARCHITECT	61–80 >80	61–80: 10,790 >80: 1931	ND	61–80: 9.7–19.3 >80: 9.7–19.6
Barhanovic et al. [124]	Abbott ARCHITECT	60–69	130 F Routine tests over two consecutive years as a part of their health checkup	9.5–19.0	11.01–19.08
Fontes et al. [137]	Roche cobas Elecsys	≥60	360 Routine tests for which evaluation of thyroid function and / or autoimmunity was not required	11.6–ND	9.0–21.9
Yeap et al. [152]	Roche cobas Elecsys	70–89	411 Apparently healthy men	10.0–23.0	12.1–20.5

cRI: conventional reference interval as established by local clinical laboratory protocols, eRI: experimental reference interval, F: female; FT4: free thyroxine, M: male, ND: not disclosed.

type-3 all influence the relationship between TSH and FT4 [164]. Misinterpretation of maternal thyroid status is a major risk factor in the care of pregnant women. For this reason, the American Thyroid Association guidelines for the diagnosis and management of thyroid disease during pregnancy and postpartum recommend using trimester-specific RIs to assess thyroid function [165].

As has been described for the general adult population, each laboratory should validate their FT4 RIs for their specific population, taking into consideration potential thyroid function determinants (such as assay methods, criteria to enroll subjects, iodine status, ethnicity, BMI, smoking, etc.).

Ho et al. established gestational age-specific RIs for serum thyroid hormones in a prospective study of 926 pregnant women from multiple ethnic groups [166]. Serum concentrations of thyroid hormones (TSH, FT4, FT3, T4, total T3), thyroid peroxidase antibody, and thyroglobulin antibody were measured using Abbott ARCHITECT immunoassays at four timepoints across the three trimesters. Although dissimilar from those for non-pregnant adults provided by the manufacturer of the assay, the study found no differences in FT4 RIs between the ethnic groups apart from at 28–32 weeks' gestation, highlighting the importance of establishing gestational age- and ethnicity-specific RIs. This finding is consistent with that described by Veltri et al., [167] but contrasts with that previously reported by Korevaar et al. in pregnant women living in an area with sufficient iodine [168]. Undoubtedly, further investigations need to be conducted in different regions of the world on racial differences in thyroid hormone parameters.

The maternal characteristic that had the greatest influence TFT was BMI. Women with obesity showed higher TSH and lower FT4 values than non-obese women [169,170]. The effect of maternal smoking was minor, although studies in the general population showed that smoking tended to be associated with lower TSH levels and with an increase in peripheral free thyroid hormones [171].

Over the years, several studies have been published with the aim of providing FT4 RIs for pregnant women [172,173] (Table 8; [162,166,172,174–196]). In most of the studies, serum FT4 showed a slight upward trend in the first trimester and decreased gradually in the second and third trimesters of pregnancy, measuring around the lower reference limit of the RI of non-pregnant women. Thus, FT4 concentrations in pregnancy can be misinterpreted as hypothyroid [197–199]. As shown by Joosen et al., FT4 levels were statistically significantly different between trimesters and postpartum ($p < 0.0001$) [172] (Figure 6). Most of the studies discussed by Joosen et al. had a retrospective, cross-sectional design that produced trimester-specific RIs, the application of which depended on reproducibility in different populations, especially when different laboratory methods were used. The impracticability and the costs of establishing a specific reference method for pregnant women make the use of pre-established RIs a very appealing approach. In a longitudinal study, a Danish group provided specific RIs for gestational age based on the use of several immunoassays [200]. They showed that the use of predetermined RIs provided a reliable option for the interpretation of TSH values, but resulted in the incorrect classification of up to 100% in the interpretation of FT4 values of pregnant women, even within populations in the same region and when the same methodological approach was used to establish the RIs. Thus, to ensure the safe care of pregnant women, the authors proposed the standardization of the Z-score of FT4 levels among laboratories to overcome these differences [200].

To overcome the influence of the concentrations of the binding proteins on FT4 immunoassays [201], some studies have evaluated serum levels of FT4 in the different trimesters of pregnancy using LC-MS/MS. Anckaert et al. compared FT4 results obtained from currently used immunoassays (Abbott ARCHITECT, Roche Diagnostics cobas and Siemens IMMULITE 2000) with those provided by equilibrium dialysis isotopic dilution (ID)-LC/MS-MS, concluding that the immunoassays produced values suitable for the clinical evaluation of thyroid function during pregnancy [87]. However, they found that the results obtained with on Abbott ARCHITECT did not show the same pattern observed using ID-LC-MS/MS or the other two immunoassays (Roche Diagnostics cobas and Siemens IMMULITE 2000) since the decline in FT4 during the latter trimesters was much less marked than for the other three methods. This observation was attributed to the increased sensitivity of the ARCHITECT assay to altered binding proteins during pregnancy. However, the main limitation of this study was the limited number of samples included in each trimester [87].

More recently, Geno et al. measured TSH and FT4, using the Roche cobas platform, in stored clinical specimens from 147 non-pregnant women of childbearing age and 580 pregnant women in the first, second, and third trimesters. A fraction of these samples was used to measure FT4 with equilibrium dialysis. A comparison of the two methods supported the use of automated FT4 immunoassays during pregnancy when the results were interpreted in the context of the appropriate

Table 8. Comparison of gestational RIs of FT4 using different platforms, in studies published from January 2012 to May 2022 including at least two different trimesters of pregnancy.

			FT4 RIs (2.5–97.5)			
Manufacturer/ analytical platform	Reference	Country	1st trimester (No. samples)	2nd trimester (No. samples)	3rd trimester (No. samples)	NP (No. samples)
Abbott	Quinn et al. [179]	Mexico	9.7–17.9 (165)	9.5–16.7 (181)	8.4–14.4 (211)	10.7–17.6 (104)
ARCHITECT System	Shen et al. [180]	China	10.9–17.7 (365)	9.3–15.2 (346)	7.9–14.1 (480)	9.2–16.9 (360)
	Fan et al. [175]	China	NS: 7.2–19.8 (140) S: 12.8–18.6 (200)	NS: 9.8–17.7 (184) S: 10.6–15.3 (200)	NS: 9.8–17.9 (123) S: 9.6–14.4 (200)	7.5–21.1†
	Liu et al. [177]	China	12.4–19.1 (312)	9.9–18.1 (304)	9.1–14.9 (331)	12.3–18.9 (150)
	Akarsu et al. [174]	Turkey	10.3–18.1 (945)	10.3–18.2 (1120)	10.3–17.9 (395)	10.7–18.2 (220)
	Ho et al. [166]	Singapore	11.4–19.5‡ (561)	10.1–15.4§ (557)	9.5–14.9| (560)	ND
	Ollero [178]	Spain	10.9–16.0 (288)	10.6–15.4 (252)	8.6–13.6 (236)	ND
	Yang et al. [181]	China	11.7–19.7 (41,634)	ND	9.1–14.4 (41,634)	ND
	Yuen et al. [182]	Hong Kong	10.5–19.4	10.5–19.4	9.5–15.3	1st T: 10.9–17.7
	Hernandez et al. [176]	Spain	10.4–16.0 (270)	8.4–12.7 (212)	8.2–12.5 (211)	9.0–19.0*
Beckman Coulter	Ekinci et al. [162]	Australia	5.9–15.6 (129)	4.9–11.3 (84)	4.4–11.2 (71)	3.8–6.0† (70)
DxI 800	Zhang et al. [193]	China	8.7–15.2 (1521)	7.1–13.6 (1102)	3.2–12.0 (120)	ND
	Liu et al. [177]	China	9.1–15.7 (312)	6.6–13.5 (304)	5.9–12.8 (331)	9.1–15.2 (150)
	Sun et al. [192]	China	9.0–15.1 (1954)	6.8–11.3 (2839)	6.7–11.4 (2056)	9.4–15.9 (646)
	Kim et al. [189]	Korea	10.8–18.4 (135)	8.8–15.6 (143)	8.6–14.5 (139)	ND
	Yuen et al. [182]	Hong Kong	8.4–16.2 (524)	7.8–14.4 (524)	6.7–11.3 (524)	1st T: 6.7–14.1
Roche Diagnostics	Moon et al. [190]	South Korea	10.7–21.2 (120)	9.1–15.7 (211)	8.4–14.6 (134)	12.1–19.3 (206)
Modular E170	Joosen et al. [172]	The Netherlands	11.7–20.0 (97)	9.3–14.2 (94)	8.1–14.9 (93)	10.8–18.2¶ (93)
cobas e 601	Sekhri et al. [191]	India	9.8–18.5 (86)	8.5–19.4 (86)	7.4–18.3 (86)	10.9–22.1 (124)
cobas e 602	Liu et al. [177]	China	13.1–22.2 (312)	9.8–18.9 (304)	8.7–15.4 (331)	13.5–20.0 (150)
Elecsys 1010	Zhang et al. [194]	China	SLRI group: 13.4–19.0** (99) CSRI group: 13.1–20.3 (957)	SLRI group: 10.5–16.2 (99) CSRI group: 9.2–17.7 (252)	SLRI group: 7.5–15.0 (99) CSRI group: 6.9–13.4 (91)	19.5–19.6 (301)
	Fan et al. [175]	China	NS: 13.4–22.5 (140) S: 13.0–20.0 (200)	NS: 10.0–17.8 (184) S: 10.4–16.0 (200)	NS: 8.7–14.8 (123) S: 9.5–14.6 (200)	12.0–22.0†
	Khalil et al. [188]	United Arab Emirates	11.7–20.4 (136)	9.3–17.2 (146)	8.7–15.3 (132)	ND
	Donovan et al. [184]	Canada	11.0–19.2 (124)	10.5–18.2 (140)	9.0–16.1 (142)	1st T: 12.1–19.6† 2nd T: 9.6–17.0† 3rd T: 8.4–15.6†
	Zhou et al. [195]	China	11.5–21.5 (1183)	9.9–18.7 (1729)	8.8–15.2 (148)	1st T: 11.8–18.4† 2nd T: 11.6–17.4† 3rd T: 9.7–15.1†
	Yuen et al. [182]	Hong Kong	11.4–24.5	10.1–22.2	9.0–17.0	1st T: 2.1–19.6†
	Hernandez et al. [176]	Spain	11.5–19.1 (270)	9.7–14.7 (212)	8.9–14.5 (211)	12.0–22.0†
	Bunch et al. [183]	USA	12.0–18.5 (453)	10.2–16.6 (479)	ND	ND
Siemens Healthineers	Duan et al. [185]	China	12.3–18.9 (963)	11–15.5 (981)	9.5–16.3 (792)	ND
ADVIA Centaur	Han et al. [186]	China	11.8–18.4 (188)	11.6–17.4 (133)	9.7–15.1 (157)	11.5–22.7†
	Zhang et al. [196]	China	13.9–26.5 (288)	12.3–19.3 (255)	11.4–19.2 (262)	13.0–22.2 (282)
	Yuen et al. [182]	Hong Kong	11.3–20.3	10.9–19.4	10.1–16.0	ND
	Huang et al. [187]	China	11.9–18.8 (8053)	11.9–18.2 (8036)	10.2–17.4 (7612)	16.1–20.0 (8646)

CSRI: cross-sectional reference interval, FT4: free thyroxine, No.: number, ND: not disclosed, NP: non-pregnant, NS: non-sequential, PP: post-partum, S: sequential, SLRI: self-sequential longitudinal reference interval, T: trimester of pregnancy, TSH: thyrotropin.

†Manufacturer Reference Intervals.

‡9–14 weeks.

§18–22 weeks.

|28–32 weeks.

¶PP.

*Compared with the CSRI group, $p < 0.05$.

**Compared with the CSRI group, $p < 0.01$.

trimester-specific RIs [202]. Hernández et al. established trimester-specific RIs for TSH and FT4 in a cohort of healthy pregnant women in Catalonia, Spain [176]. This prospective, observational study was conducted on 332 healthy pregnant women from the first trimester to delivery. FT4 was measured using two immunoassays (Abbott ARCHITECT and Roche Diagnostics cobas) in the three trimesters, and by isotopic dilution LC-MS/MS after an ultrafiltration separation step in the first trimester. Results showed that although correlated, FT4 results measured by LC-MS/MS and the two immunoassays were not interchangeable.

In summary, several studies have been published in recent years with the aim of providing RIs for thyroid hormones in pregnant women [203]. Although the conclusions reached were not always consistent, overall,

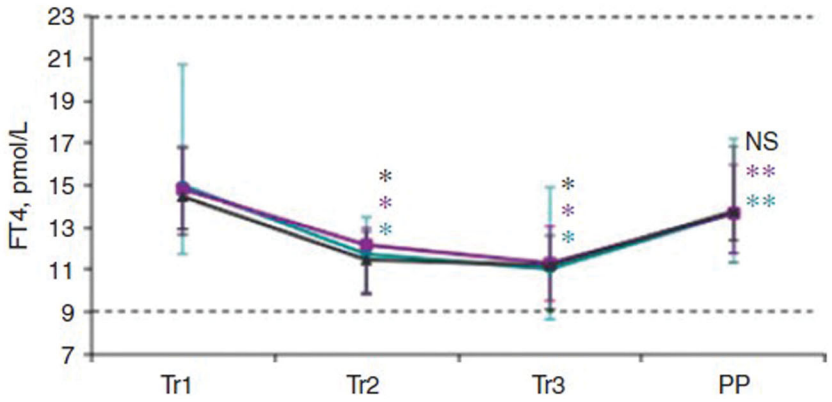

Figure 6. FT4 concentrations during pregnancy and post-partum. Error bars represent the 5th and 95th percentiles per trimester. Dashed lines represent non-pregnant lower and upper reference limits. (Green, 1st tertile; violet, 2nd tertile; grey, 3rd tertile; $^{*}p < 0.0001$ compared with Tr1; $^{**}p < 0.01$ compared with Tr1; NS, not statistically significant compared with Tr1m; FT4, free thyroxine; NS, non-significant; Tr1, trimester 1; Tr2, trimester 2; Tr3, trimester 3; TSH, thyroid-stimulating hormone; PP, post-partum).

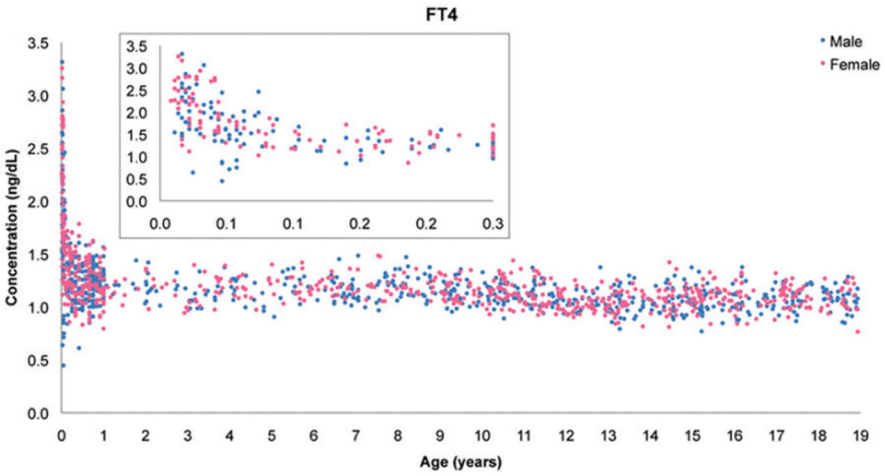

Figure 7. Age-dependent scatter plot of FT4 concentration, stratified by sex. (FT4, free thyroxine).

nearly all of the studies found that thyroid status changed with gestational age. However, while the temporal trend of FT4 levels was similar between different populations, the actual hormone values were not; they showed significant intra- and inter-method discrepancies (Table 7). Although the importance of specific-trimester RIs has been widely emphasized, the importance of the method used to determine such RIs has been neglected [173]. In the future, it is expected that methods for FT4 will be compared to and calibrated with LC-MS/MS, the candidate gold standard method [200].

FT4 reference intervals in the neonatal, pediatric, and adolescent age groups

Thyroid hormones are essential for normal child growth and development [204]. Hypothyroidism in children is associated with intellectual disability, short saturation, delayed skeletal maturation, and puberty [205]. Conversely, hyperthyroidism in children is associated with hyperactivity, irritability, poor schooling performance, and other abnormalities [206]. Thus, prompt treatment of newborns with abnormal thyroid function ensures normal physical and mental development [207].

Guidelines of the European Thyroid Association for the management of subclinical hypothyroidism in children recommend the use of age-related RIs [165,208]. However, accurate RIs according to age and sex are not yet readily available in pediatric laboratory medicine [209].

Over the past decade, national and international initiatives have begun to fill this gap. In 2012, the CALIPER study was the first to publish specific RIs for several common biochemical markers in Canadian children (Figure 7) [210,211]. To date, the CALIPER study has established age- and sex-specific RIs for thyroid hormones on the Abbott ARCHITECT c8000 [210,211], Beckman Coulter [212], Roche Diagnostics cobas [150], Ortho VITROS [213], and Siemens Healthineers Atellica

Table 9. Comparison of RIs of FT4 for children aged 1–18 years using different platforms, in studies published from January 2012 to May 2022.

Manufacturer/ Analytical platform	Reference	Country	Age group	Sample size (No.)	FT4 RIs (2.5–97.5 percentile)
Abbott	Aldrimer et al. [215]	Sweden	0.5–12 y	471	10.8–16.4
ARCHITECT System			13–18 y	215	10.2–15.5
	Radicioni et al. [219]	Italy	6.2–12.1 y	72	13.1–20.6
			9.6–17.9 y	368	10.9–19.1
	Bailey et al. [211]	Canada	1–<19 y	952	11.4–17.6
	Bokulic et al. [218]	Croatia	1–<19 y	241	10.5–15.9
	Argente del Castillo et al. [217]	Spain	1–9 y	531	11.1–15.4
			10–13 y	577	9.9–14.9
			14–15 y	473	9.9–14.2
Beckman Coulter	Romero et al. [220]	Mexico	13–18 m	47	14.0–35.1
Dxl 800			19–23 m	48	16.0–33.0
			2–<3 y	73	15.6–30.4
			3–<6 y	62	14.0–28.2
	Karbasy et al. [212]	Canada	3–<19 y	455	7.9–13.6
	Adeli et al. [209]	Canada	1–<19 y	982	13.0–21.0
Roche Diagnostics	Iwaku et al. [222]	Japan	4–6 y	43	14.4–21.5
Modular E170			7–8 y	39	13.8–20.7
cobas e 601			9–10 y	51	12.4–20.6
cobas e 602			11–12 y	61	13.1–19.6
Elecsys 1010			13–14 y	72	19.4–19.6
			15 y	50	12.2–19.7
	Bohn et al. [150]	Canada	1–<19 y	982	13.0–21.0
	Gunapalasingham et al. [221]	Denmark	6–18.9	2407 (976 M; 1435 F)	ND
			6–9.9	866 (395 M; 471 F)	10.9–18.4
			10.0–14.9	1021 (605 M; 416 F)	10.0–17.8
			15.0–18.9	524 (165 M; 359 F)	10.2–17.8
Siemens Healthineers	Strich et al. [224]	Israel	1–5 y	2722	11.0–19.0
Advia Centaur			6–10 y	3452	11.3–18.7
			11–14 y	3429	10.5–17.9
			15–18 y	1019	10.4–18.0
	Loh et al. [223]	Singapore	7–10 y	NS	10.9–20.6
			10.1–12 y		10.2–20.1
Siemens Healthineers Atellica	Bohn et al. [214]	Canada	1–<19 y	810	13.4–21.1

FT4: free thyroxine, m: months, No.: number, ND: not disclosed, NS: not specified, RI: reference interval, y: years.

systems [214]. Experience from the CALIPER study has encouraged other research groups to estimate FT4 RIs in children and adolescents [215,216]. As seen in the adult population, published data highlighted significant discrepancies between studies, suggesting that FT4 RIs are dependent on several factors, including the method used and the population enrolled (Table 9; 150,209,211,212,214,215,217–224) [218, 225]. Much of the scientific literature published to date agrees that there are no significant differences between serum FT4 concentrations in the distinct age groups [150,211,221]. However, some studies do describe statistically significant differences in FT4 concentrations between the pediatric and adolescent age groups and/or between males and females [148,217,226].

Studies have also reported that thyroid hormones show large variations in concentration during the first year of life [150,211,227]. In particular, the 2.5th and 97.5th percentiles of FT4 are higher in the first days of life and have high biological variation, suggesting that a larger RCV is essential for serial results to be considered significantly different [223,228–230]. After the first year of life, FT4 values tend to decrease with increasing age for both sexes. Although each study considered different age groups, FT4 levels generally remained constant throughout childhood, declining again in the prepuberty period (9–15 years) when several important sex-dependent changes occur to the morphology and physiology of the adolescent [150,215–217,221,225,226,231,232]. The physiological processes underlying the changes observed in thyroid hormones in the time span from birth to adulthood have yet to be fully delineated. At birth, a newborn quickly adapts to extrauterine life by developing a state of relative overactivity of the thyroid gland [233]. A sudden burst of thyrotropin-releasing hormone and TSH release, reaching up to 70–100 mIU/L within 30 min of birth, leads to a two- to six-fold increase in circulating T4 and T3 concentrations. TSH significantly decreases to within normal infant concentrations in the first 3–5 days of life, while FT3 and FT4 serum levels remain elevated for several days and act on tissues. Thus, the interpretation of RIs of TSH and FT4 for newborns must take into account the gestational age and postnatal age up to one month old (Table 10; [209,211,212,218,220,223,224,229,234–237]) [229, 237, 238]. Thyroid hormone concentrations then decrease slightly during childhood and adolescence. There is also a progressive decrease in thyroid T4 production, iodine turnover, and absorption with age, which produces an

Table 10. Comparison of RIs of FT4 for infants aged 0–12 months using different platforms, in studies published from January 2012 to May 2022.

Manufacturer/ analytical platform	Reference	Country	Age group	Sample size (No.)	FT4 Reference Intervals (2.5–97.5 percentile)
Abbott ARCHITECT System	Bailey et al. [211]	Canada	5–14 d	66	13.5–41.3
			15–29 d	55	8.7–32.5
			30 d–11.9 m	270	11.4–21.9
	Bokulic et al. [218]	Croatia	2–15 d	68	11.8–28.0
			15 d–<12 m	54	11.3–18.9
Beckman Coulter DxI 800	Romero et al. [220]	Mexico	1 d–<1 m	47 (25 M; 22 F)	8.6–34.6 (8.4–30.6 M; 11.3–35.4 F)
			1 m–<6 m	76 (42 M; 34 F)	8.5–15.6 (7.8–15.4 M; 8.6–17.2 F)
			6–<12 m	52 (30 M; 22 F)	8.2–14.4 (8.1–12.1 M; 8.5–14.7 F)
	Karbasy et al. [212]	Canada	0–<20 d	80 (40 M; 40 F)	17.4–57.7
			20 d–3 y		9.5–17.8
	Adeli et al. [209]	Canada	0–<20 d	40	17.4–57.7[†]
			20 d–3 y	215	9.52–17.8[†]
	Jayasuriya et al. [229]	Australia	0–24 h	25	15.3–43.6
			25–48 h	51	14.7–53.2
			49–72 h	71	16.5–45.5
			73–96 h	86	17.8–39.4
			97–120 h	63	15.3–32.1
			121–144 h	32	14.5–32.3
			145–168 h	62	13.9–30.9
	Aktas et al. [234]	Turkey	4–7 d	482	18.66 ± 4.24[‡]
			8–14 d	131	16.73 ± 2.57
			15–22 d	57	14.93 ± 2.83
			23–30 d	87	14.28 ± 3.86
	Wong et al. [237]	Malaysia	14–21 d	513	11.1–21.0
			22–30 d	66	10.1–19.6
Roche Diagnostics Modular E170 cobas e 601 cobas e 602 Elecsys 1010	Mutlu et al. [238]	Turkey	1 d	29	15.4–33.6
			3 d	39	15.4–42.5
			5 d	28	13.6–29.6
			7 d	30	14.5–34.6
			10 d	82	15.2–32.1
			14 d	22	14.5–28.7
			28 d	19	15.8–24.9
	Omuse et al. [236]	Kenya	1–7 d	552	13.6–34.8
			8–15 d	145	13.5–30.2
			15–20 d	465	14.2–24.8
			23–30 d	167	13.3–23.4
Siemens Healthineers Advia Centaur	Loh et al. [223]	Singapore	0–1.0 w	NS	19.9–46.6
			1.1–2.0 w		17.2–33.1
			2.1–3.0 w		15.0–25.9
			3.1–4.0 w		13.22–21.8
			1.1–2.0 m		11.3–21.3
	Strich et al. [224]	Israel	1 d–<1 m	47	12.4–27.4
			1 m–<2 m	58	12.4–21.8
			2 m–<12 m	317	10.8–19.5

[†]Mean ± SD.
[‡]2SD from mean.
d: days; F: females; FT4: free thyroxine; h: hours; M: males; m: months; No.: number; ND: not disclosed; NS: not specified; RI: reference interval; SD: standard deviation; w: weeks.

overall progressive decrease in thyroid function. The concomitant drop in TSH during this period suggests that it is the primary mediator of these effects. Serum TBG concentrations increase up to age 5 years; this increase contributes to the gradual dissociation between FT4 and T4 [233,239]. Subsequently, the TBG concentration, which decreases between 15 and 16 years of age, results in a gradual decrease in serum concentrations of total T3 and T4 [216,225,240].

Similar to the factors identified for adults, the determination of FT4 RIs in children and adolescents depends on the type of assay, sample size, age, sex, ethnicity, geographical region, time of venipuncture, fasting or non-fasting condition, BMI, and other anthropometric characteristics that have been identified as determinants of thyroid function. However, in this population, some of these determinants deserve to be further explored. First, the sample size must be considered. Many studies lack sufficient numbers of participants to calculate appropriate RIs by age range. In fact, although 120 participants have been indicated as the required number to calculate the 2.5th and 97.5th percentiles with a confidence interval of 90%, in the case of neonatal and pediatric populations, which are characterized by high inter-individual variability, the number must rise to at least 400 participants to ensure

reliable RIs [216]. To achieve this goal, recent studies have used an indirect method; however, such an approach carries the risk of including in the study individuals with thyroid dysfunction whose influence on the results is not quantifiable, although it is generally believed to be minimal [217].

Another important aspect is the age of the child/adolescent. Most of the studies on FT4 RIs use age expressed in years, while pediatricians use Tanner stages to better characterize the physical development of children and adolescents. Furthermore, in studies including the adolescent phase, the assessment of pubertal status was limited to the pre-pubertal/post-pubertal period and data did not allow a detailed assessment of the relationships between pubertal stages and thyroid function or associations with other parameters such as growth rate or circulating levels of sex hormones or insulin-like growth factor 1 [231]. The role of weight, height, and body composition in determining serum concentrations of TSH and FT4 is also still debated [148]. Finally, the literature on childhood FT4 RIs often does not consider the timing of venipuncture or the state of fasting, making a comparison between studies even more difficult [223]. Fasting and feeding status are not easy to investigate as children often do not eat at set times and may consume snacks during the visit [216].

In conclusion, heterogeneity between studies published so far highlights the difficulty in harmonizing pediatric thyroid RIs and demonstrates the need for standardized RI for methodologies used in this area [217]. Undoubtedly, the identification of thyroid determinants and the quantification of their effects can help with the interpretation of TFT. Future efforts should focus on generating evidence-based recommendations for defining abnormal thyroid function and RIs in children.

FT4 reference intervals in patients on levothyroxine therapy

When FT4 RIs is determined, patients with hypothyroidism on replacement therapy with levothyroxine should be considered, as these individuals undergo very frequent TSH and FT4 measurements. FT4 concentrations showed a clear shift to the right of the distribution curve compared with control subjects, with a portion of these patients (10%) having FT4 concentrations higher than the manufacturer's upper reference limit and normal TSH concentration values [48,241,242]. Notably, patient-related factors such as changes in thyroid hormone receptor sensitivity, age, and sex add further complexity to the challenge of establishing specific RIs for patients treated with levothyroxine. In practice, it is preferable to monitor TSH levels in this population and to use FT4 measurement only in patients with central hypothyroidism. Conclusions about the need for specific serum FT4 RIs were reported by a recent study on central hypothyroidism in children on T4 replacement therapy [243].

Reference change value (RCV)

RIs are of limited use in evaluating serial results obtained on an individual [244]. In fact, each individual has a range of values that covers only part of the population-based RIs. Consequently, individuals can have significant changes in results even within the RI: such changes are often considered insignificant and therefore disregarded by both laboratory professionals and physicians. Furthermore, the results can shift from within the RI to outside the RI (and vice versa) with no clinical significance: laboratories conventionally flag results outside the RIs, likely initiating unnecessary follow-up activity [245]. The RCV, or critical difference, is a more useful means of evaluating differences in serial test results. The basis for this simple tool is that, for a change to be significant, the difference between two measurements must be greater than the inherent variation, which is explained by preanalytical variation (CV_P), inter-assay analytical variation (CV_A), and within-subject (CV_I) biological variation [244]. A normal distribution of analytical and biological variability is usually assumed, and for each Z score value, corresponding to a probability that the variation is not due to chance, the difference between the two test results obtained from the same subject can be calculated with the following formula:

$$\mathrm{RCV} = 2^{1/2} \cdot Z \cdot \left(CV_P^2 + CV_A^2 + CV_I^2\right)^{1/2}.$$

For many tests, including FT4, the change is smaller than that required for the result to fall outside the RI. In addition, in monitoring, the primary interest is focused on unidirectional change, such as a drop in FT4 concentrations following the treatment of Graves' disease [61]. Data has been published on intra-subject biological variability for most thyroid-related tests [246–249]. Recent data from the European Biological Variation Study (EuBIVAS) population, which was composed of 91 healthy individuals from six European laboratories, 21–69 years old, showed a CV_I of 4.8% for FT4 with no significant differences between men and women [250]. EuBIVAS provides lower biological variability estimates than those previously published. This implies not only more stringent analytical performance but also smaller RCVs that would identify minor changes

as being within the expected variation caused by analytical and biological variation. If the FT4 analytical and biological variation estimates derived from EuBIVAS were used, the RCV (95% probability for a significant unidirectional change) was 13.1%. In clinical practice, this means that, for an adult patient with an FT4 value of 16.7 pmol/L, an increase to 18.8 pmol/L could be explained simply by biological and analytical variations. However, it is important to bear in mind that results within the RCV do not rule out the clinical importance of such a change.

Another study that included 19 healthy volunteers (8 male and 11 female) with blood drawn every week for 5 consecutive weeks showed that the CV_A, CV_I, and between-subject biological variation (CV_G) were 3.6%, 4.6%, and 10.8% for FT4, respectively. The index of individuality (II) for all parameters was between 0.2 and 0.7. The percentage above which the change between the two measures is truly significant (RCV) was 16.2% for FT4 [251]. Data on biological variation over 24 h was obtained from 31 healthy subjects at 0000, 0400, 0800, 1200, 1600, and 2000 h; the CV_I and CV_G estimates were 3.57% and 8.03% for FT4. The IIs of all the biomarkers (except TSH) were $\leq$0.63. Males had lower CV_Is and IIs than females [252]. The use of RCV in a specific patient should allow early detection of significant changes with respect to the population RI and is likely to be more reliable than a physician's intuition. In addition, once a large number of serial measurements for a patient are available, a variety of statistical tests can be used to discern significant changes, assuming that the dosage has been constant and that the patient has not been exposed to diseases or drugs that may affect the biological variability of the analyte in question [61]. Despite all these advantages, the RCV approach is currently under-used. In part, this is because laboratory tests are increasingly dissociated from clinical practice, making it difficult to match patients and results serially. Additionally, it is not uncommon for patients to be tested for a particular analyte with different assays over time. The persistent lack of harmonization among different assays effectively precludes the use of RCV in such patients. Reliable biological variability characteristics, and especially RCV, can facilitate the interpretation of consecutive TFT in an individual and therefore have the potential to support clinical decisions regarding thyroid diseases efficiently [250].

Analytical interferences in immunoassays

Thyroid dysfunctions are commonly diagnosed by evaluating thyroid-specific symptoms along with laboratory measurements of TSH, FT3, and FT4 in blood. Although these measurements are routine examinations, the analytical procedures for determining TSH, FT3, and FT4 pose major challenges for laboratory diagnostics due to multiple interference factors. Falsely increased or decreased thyroid hormone measurements that are caused by interference factors in immunoassays may result in a considerable number of misinterpretations of laboratory findings [62,253]. Numerous factors can interfere with measurements of TSH, FT3, or FT4 in immunoassays. Macro molecules (frequency $<$1/100) [254], interfering antibodies (frequency $<$1.1/100) [255], various amino acid or glycosylation variants (frequency $<$1/100,000) [256,257], and exogenous intake of biotin (see below) may falsify TSH results.

FT4 interferences can be categorized into (a) interferences in the analytical procedure of the immunoassay or LC-MS/MS, (b) matrix effects, (c) high or low levels of thyroid hormone distributor proteins (THDPs) – TBG, TTR, and albumin, (d) genetic variants of THDPs, (e) modifications in the binding ability of thyroid hormones by THDP-triggered drugs, (f) interfering antibodies, and (g) biotin intake. While a similar list of interferences affects the measurement of FT3, this review focuses on FT4 interferences.

Analytical procedures

In the measurement of free thyroid hormones, various interferences may lead to variations in the final measured value. Thus, the validation of FT4 immunoassays without physical separation of the free hormone by equilibrium dialysis- or ultrafiltration-coupled FT4 immunoassays or LC-MS/MS methods has identified an overestimation of FT4 at low concentrations (see Measurement of free thyroxine) [75].

Serum matrix

Constituents of the serum matrix influence the measurement of FT4 indirectly. Increased values of, for example, hemoglobin, bilirubin, or lipids may reduce the measurement signal and cause falsely increased hormone levels in competitive immunoassays. The manufacturer's package insert for each assay specifies limits of acceptance for the levels of these potentially interfering substances, which differ depending on which FT4 method is used. In cases in which contamination of the sample is strongly suspected, blood sampling must be repeated [258].

Thyroid hormone distributor proteins

Over the last 20 years, a number of studies have demonstrated the dependence of FT4 measurements on the level of THDPs. Such interaction is expected theoretically only for the measurement of T4 by immunoassays and not primarily for FT4. However, FT4 immunoassay values without physical separation of the free fraction also revealed a method-dependent association with THDP levels, especially in cases with distinctly increased TBG, albumin, and immunoglobulin G levels [16,259].

Genetic variants of thyroid hormone distributor proteins

Patients with genetic variants of THDP are mostly euthyroid, but because of reduced THDP-binding ability, the values of T4 or FT4 are frequently falsely elevated. This is true primarily in cases of familial dysalbuminemic hyperthyroxinemia (prevalence 0.01–1.8%, with the highest values in the Hispanic population) [260]. If this disease is suspected, the use of a physical separation method-coupled immunoassay or LC-MS/MS method to measure unbiased FT4 is necessary. TBG variants are usually associated, not with thyroid diseases, but with abnormal T4 (with TBG deficiency, T4 is low; with TBG excess, T4 is high) and normal free thyroid hormone values [261]. Mutations of TTR present an increased or decreased binding affinity for T4. While decreased TTR affinity to T4 has no effect on serum levels of thyroid hormone, increased TTR affinity leads to mildly increased T4 and normal FT4 levels if a physical separation method-coupled immunoassay or an LC-MS/MS method is used to quantify FT4 [261].

Drugs

Some drugs (e.g. aspirin, furosemide, phenytoin) may displace the equilibrium between thyroid hormones and THDP; others may increase (e.g. estrogen, fluorouracil, tamoxifen) or inhibit (e.g. androgens, glucocorticoids, nicotinic acid) the synthesis of TBG, leading to questionable FT4 results [262]. Details about the characteristics and degree of potential drug interactions are described in the manufacturer's package insert for each FT4 assay. A specific issue is the administration of heparin, which can cause an artefactual elevation in FT4 by displacing thyroid hormones from THPD *via* promptly generated non-esterified fatty acids, especially when FT4 is measured by equilibrium dialysis [263].

Heterophilic antibodies

Heterophilic antibodies (HAbs) are the currently used imprecise nomenclature for nonspecific endogenous antibodies against animal immunoglobulins used for analyte-binding reactions in immunoassays. This includes human anti-mouse antibodies, human anti-animal antibodies, and Fc-region binding rheumatoid factor-associated immunoglobulin. The prevalence of HAbs in human serum has been cited as 0.08% and 6% [255,264].

As the presence of HAbs in FT4 immunoassays usually leads to an inhibition of binding signals, the consequences of this effect are spuriously increased FT4 values. Although proteins with higher molecular weight (e.g. TSH) are more susceptible to HAbs than small thyroid hormone molecules, case reports also describe this interference for FT4 or T4 [264–268].

Anti-streptavidin antibodies

Anti-streptavidin antibodies may lead to decreased TSH and increased FT4 values in assays that use biotin-streptavidin as an immobilizing system to enhance the binding capacity and subsequently the sensitivity and measurement range of immunoassay platforms. This finding is quite similar to the interference due to biotin intake (see below) and can be differentiated only by detecting the immunoglobulin G or M antibody characteristics (e.g. *via* molecular weight by size exclusion chromatography) [269,270].

Anti-ruthenium interference

Anti-ruthenium interference can be observed in immunoassays developed by the manufacturer, Roche Diagnostics, which uses this rare transition metal as a label for the generation of electrochemiluminescence. Although falsely decreased TSH and/or elevated FT4 values were reported in most published cases, falsely decreased FT4 values were also detected, demonstrating the heterogeneity of this antibody [62,271].

Anti-thyroid hormone antibodies

Anti-thyroid hormone antibodies (THAb) are present as immunoglobulin G and M isotypes with a serum prevalence of $<2\%$ [272] in healthy subjects and up to 71% [273] in patients with autoimmune thyroid diseases. This prevalence depends on not only the cohort investigated but also the sensitivity of the antibody detection method. The potential interfering effects of THAbs on FT3 and FT4 measurements are driven by the characteristics of the *binding affinity* (*K*) and *binding capacity* (BK) of the endogenous compared with the analytical immunoassay-associated THAb. It is possible for the FT4

values to be increased only if the *K* and BK of the analytical antibody are distinctly lower than those of the endogenous THAb. Although one-step free hormone assays are susceptible mainly to THAb interference, falsely increased FT3 and FT4 values due to such interferences do not always occur [274]. As only around 9% [273] of patients with autoimmune thyroid disease revealed falsely elevated FT4 values (depending on the method), the measurement of THAb is usually not performed in routine diagnostics.

Biotin

Biotin-labeled antigens (also known as vitamins H, B7, and B8) or antibodies are part of the biotin-streptavidin system that is used to improve immobilization and, therefore, the signal capacity of approximately half of the immunoassay methods. Exogenously applied biotin may reduce the measurement signals and lead to increased antigen levels in competitive assays and reduced antigen levels in sandwich tests that use this system. Biotin also serves as a nutrition supplement ($<$20 mg) for alopecia and for the health of hair, nails, and skin. High doses (100–300 mg) of biotin have been used to treat biotinidase deficiency and multiple sclerosis. Consequently, a biotin dose of 10 mg/day over 8 days was sufficient to significantly reduce TSH levels (sandwich assay), by up to 34%, and increase FT4 levels (competitive assay), by up to 13% in the Roche Diagnostics cobas system, but no changes were observed for the Abbott ARCHITECT platform 2 h after ingestion [275]; despite this, the measured levels were still within the RI. Moreover, biotin levels were negatively correlated with TSH (-0.29, $r = 0.01$) and positively correlated with FT4 values ($r = 0.3$, $p < 0.01$). Spuriously increased TSH and FT4 values from biotin intake have also been found in other studies, for example, when the platforms Dimension Vista (TSH and FT4, Siemens Healthineers, Tarrytown, NY, USA), DxI (Beckman Coulter, FT4), and VITROS (Ortho, TSH) were used [276]. As most manufacturers have enhanced the tolerance of their assays to biotin today, this effect is mostly theoretical unless patients take very large doses of biotin (e.g. for multiple sclerosis) and have concomitant renal failure [277,278]. Nevertheless, awareness of the risks of potential interactions with exogenous biotin is important for endocrinologists.

Examples of analytical interference

Despite best efforts, there are many examples in the literature where assay interferences and subsequent misinterpretation of thyroid hormone levels have led to misdiagnosis, incorrect treatment, and potential harm to patients [62]. One example concerns a pregnant woman who was diagnosed with hyperthyroidism based on her clinical symptoms (palpitations, dyspnea, and tachycardia) and a discordant thyroid hormone profile (elevated FT3, elevated FT4, and TSH within the RI), which persisted throughout her pregnancy [279]. Her daughter also had a similar thyroid hormone profile for over a month after birth without displaying clinical signs of congenital hyperthyroidism. It was hypothesized that the abnormal thyroid hormone profiles observed were due to THAb interference, and further analytical investigations supported this by demonstrating the presence of anti-T3 and anti-T4 in the serum of the woman and her daughter. In this case, the woman was incorrectly treated with methimazole throughout her pregnancy, which highlights the importance of accurate understanding and interpretation of assay values.

Ricci et al. reported six confirmed cases of anti-streptavidin antibody assay interaction in which the patients had discordant thyroid hormone profiles without clinical symptoms [280]. All patients had TSH levels within the RI but elevated total T3 and T4; some patients also had abnormal levels of FT4. On further investigation, all patients had anti-streptavidin antibodies that were believed to have interfered with the assays used (cobas 8000 e 801 module for T3 and T4, and cobas 6000 e 601 module for FT4). These findings highlight the importance of accurate interpretation of assay results and regular communication between the laboratories carrying out the tests and clinicians.

Detection and elimination of interferences

In cases with increased FT4 and normal/increased TSH that lack a clinical correlate for thyroid disease, as a first step, measurements should be repeated with the same platform. If the results are confirmed, repeating the test with a different immunoassay platform may be helpful as most analytical interferences are strongly dependent on the analytical procedure. Thus, for example, interferences are reduced if an equilibrium dialysis- or ultrafiltration-coupled FT4 immunoassay or a clinically validated LC-MS/MS method for FT4 is used instead of a simple FT4 immunoassay. A significant difference in FT4 between the two platforms and a normalized value indicates interference in the measurement on the primary platform. If the secondary platform reveals a comparable value, the two platforms may suffer from the same analytical issue. In such cases, the measurement

of T4 may help, but only if the levels of THDP are within the RI [264]. Indications of potential high or low levels of THDP [281], genetic variants of THDP (largely an exclusionary diagnosis), or FT4-interfering drugs must be derived from the patient's medical history. Such interferences can be only roughly estimated, but, if possible, intake of the drug may be discontinued for a short time period.

Interferences in the measurement of TSH or FT4 by interfering antibodies may have already been revealed by the above-mentioned change in the assay methodology. The use of commercially available blocking tubes, such as HBR-1 (Scantibodies), TRU (Meridian), HBT50 (Skybio), RF Absorbent (Serion), Rapid Serum Tube (BD), PolyMAK-33 (Roche Diagnostics), and HeteroBlock (Omega Biologicals) [282,283], can detect spuriously increased values of an analyte. However, the detection rate for the disclosure of such interferences is often far from 100% [264]. Polyethylene glycol precipitation at a final concentration of 12.5% is an efficient tool to suppress the binding of interfering antibodies in the blood [282]. Moreover, Protein A or G binding or size-exclusion chromatography depletes interfering immunoglobulins and enables TSH or FT4 measurements without the effect of interferences [266]. Interfering antibodies are blocked in immunoassays by the manufacturer addition of animal immunoglobulins into the assay buffer. However, to date, no immunoassay is known to be free from any risk of interference by nonspecific interfering antibodies. Specific interfering antibody levels directed against streptavidin, ruthenium, or thyroid hormones can be detected by in-house assays. Moreover, the laboratory may use streptavidin-coated microparticles to adsorb biotin or anti-streptavidin antibodies in a direct way [284]. In the case of ruthenium, Roche Diagnostics developed an in-house method for the measurement of antibodies. Unfortunately, laboratories that provide THAb measurements as a diagnostic service are usually not available. Commercially available THAb kits are not expected to be available for many years. Accordingly, methods such as equilibrium dialysis, LC-MS/MS, or immunoassays with previous antibody separation methods (e.g. polyethylene glycol precipitation, treatment with protein A/protein G sepharose beads, or size exclusion chromatography purification) that are not susceptible to interference by thyroid hormone autoantibodies must be used when interference is suspected in the measurement of FT4 [264,285].

We have described a variety of interferences that may affect the measurement of FT4 as well as TSH and FT3 assays. Increased FT4 serum levels in euthyroid patients without clinical symptoms are almost always due to measurement interferences rather than thyroid axis disease. If such interferences are identified, they can usually be overcome by applying a variety of analytical tools.

Conclusions

Thyroid dysfunction is among the most common endocrine disorders and accurate biochemical testing is needed to confirm or rule out a diagnosis. Notably, true hyper- and hypothyroidism in the setting of a normal TSH are highly unlikely, making the assessment of FT4 levels inappropriate in most cases. However, FT4 measurement is integral in both the diagnosis and management of relevant central dysfunctions (central hypothyroidism and central hyperthyroidism) as well as in monitoring therapy in hyperthyroid patients treated with anti-thyroid drugs or radioiodine. In such settings, accurate FT4 quantification is required. Significant progress has been made in the standardization of procedures for FT4 testing, but technical and implementational challenges, including the establishment of clinical decision limits in different patient populations and education of all stakeholders, remain. Accordingly, different assays and reference values cannot be interchanged. Two-way communication between laboratories and clinical specialists is pivotal to properly select a reliable FT4 assay, establish RIs, approaching discordant results, and monitor the analytical and clinical performance of this method over time.

Acknowledgments

Editorial assistance was provided by Erin Slobodian, BA, and Anna King, Ph.D., of Ashfield MedComms (Macclesfield, UK), an Inizio company; this was funded by the Clinic for Nuclear Medicine and Molecular Imaging, Imaging Institute of Southern Switzerland, Ente Ospedaliero Cantonale, Bellinzona, Switzerland.

Disclosure statement

The authors report no declarations of interest.

Funding

The author(s) reported there is no funding associated with the work featured in this article.

ORCID

Federica D'Aurizio http://orcid.org/0000-0003-2311-859X
Damien Gruson http://orcid.org/0000-0001-5987-4376
Petra Petranović Ovčariček http://orcid.org/0000-0003-1946-2128
Luca Giovanella http://orcid.org/0000-0003-0230-0974

References

[1] Sheehan MT. Biochemical testing of the thyroid: TSH is the best and, oftentimes, only test needed – a review for primary care. Clin Med Res. 2016;14(2):83–92.

[2] Bauer DC, Brown AN. Sensitive thyrotropin and free thyroxine testing in outpatients. Are both necessary? Arch Intern Med. 1996;156(20):2333–2337.

[3] de los Santos ET, Starich GH, Mazzaferri EL. Sensitivity, specificity, and cost-effectiveness of the sensitive thyrotropin assay in the diagnosis of thyroid disease in ambulatory patients. Arch Intern Med. 1989;149(3):526–532.

[4] Thienpont LM, Van Uytfanghe K, Poppe K, et al. Determination of free thyroid hormones. Best Pract Res Clin Endocrinol Metab. 2013;27(5):689–700.

[5] Steele BW, Wang E, Klee GG, et al. Analytic bias of thyroid function tests: analysis of a college of American pathologists fresh frozen serum pool by 3900 clinical laboratories. Arch Pathol Lab Med. 2005;129(3):310–317.

[6] Thienpont LM, Beastall G, Christofides ND, et al. Measurement of free thyroxine in laboratory medicine–proposal of measurand definition. Clin Chem Lab Med. 2007;45(4):563–564.

[7] Thienpont LM, Beastall G, Christofides ND, et al. Proposal of a candidate international conventional reference measurement procedure for free thyroxine in serum. Clin Chem Lab Med. 2007;45(7):934–936.

[8] Thienpont LM, Van Uytfanghe K, Beastall G, et al. Report of the IFCC Working Group for Standardization of Thyroid Function Tests; part 2: free thyroxine and free triiodothyronine. Clin Chem. 2010;56(6):912–920.

[9] Baker CH, Morris JC. The sodium-iodide symporter. Curr Drug Targets Immune Endocr Metabol Disord. 2004;4(3):167–174.

[10] De La Vieja A, Dohán O, Levy O, et al. Molecular analysis of the sodium/iodide symporter: impact on thyroid and extrathyroid pathophysiology. Physiol Rev. 2000;80(3):1083–1105.

[11] Dohán O, De la Vieja A, Paroder V, et al. The sodium/iodide symporter (NIS): characterization, regulation, and medical significance. Endocr Rev. 2003;24(1):48–77.

[12] Darrouzet E, Lindenthal S, Marcellin D, et al. The sodium/iodide symporter: state of the art of its molecular characterization. Biochim Biophys Acta. 2014;1838(1 Pt B):244–253.

[13] Portulano C, Paroder-Belenitsky M, Carrasco N. The Na+/I- symporter (NIS): mechanism and medical impact. Endocr Rev. 2014;35(1):106–149.

[14] Carvalho DP, Dupuy C. Thyroid hormone biosynthesis and release. Mol Cell Endocrinol. 2017;458:6–15.

[15] Ulianich L, Suzuki K, Mori A, et al. Follicular thyroglobulin (TG) suppression of thyroid-restricted genes involves the apical membrane asialoglycoprotein receptor and TG phosphorylation. J Biol Chem. 1999;274(35):25099–25107.

[16] van Deventer HE, Mendu DR, Remaley AT, et al. Inverse log-linear relationship between thyroid-stimulating hormone and free thyroxine measured by direct analog immunoassay and tandem mass spectrometry. Clin Chem. 2011;57(1):122–127.

[17] Hoermann R, Midgley JE, Larisch R, et al. Homeostatic control of the thyroid-pituitary axis: perspectives for diagnosis and treatment. Front Endocrinol. 2015;6:177.

[18] Calebiro D, Nikolaev VO, Gagliani MC, et al. Persistent cAMP-signals triggered by internalized G-protein-coupled receptors. PLoS Biol. 2009;7(8):e1000172.

[19] Kronenberg HM, Melmed S, Larsen PR, et al. Principles of endocrinology. In Melmed S, Polonsky KS, Larsen PR eds. Williams textbook of endocrinology. Philadelphia: Elsevier Saunders; 2011. p. 3–12.

[20] Larsen PR, Zavacki AM. The role of the iodothyronine deiodinases in the physiology and pathophysiology of thyroid hormone action. Eur Thyroid J. 2012;1(4):232–242.

[21] Pappa T, Ferrara AM, Refetoff S. Inherited defects of thyroxine-binding proteins. Best Pract Res Clin Endocrinol Metab. 2015;29(5):735–747.

[22] Larsen PR. Thyroid-pituitary interaction: feedback regulation of thyrotropin secretion by thyroid hormones. N Engl J Med. 1982;306(1):23–32.

[23] Ross DS, Burch HB, Cooper DS, et al. 2016 American Thyroid Association guidelines for diagnosis and management of hyperthyroidism and other causes of thyrotoxicosis. Thyroid. 2016;26(10):1343–1421.

[24] Jonklaas J, Bianco AC, Bauer AJ, et al. Guidelines for the treatment of hypothyroidism: prepared by the American Thyroid Association Task Force on Thyroid Hormone Replacement. Thyroid. 2014;24(12):1670–1751.

[25] Garber JR, Cobin RH, Gharib H, et al. Clinical practice guidelines for hypothyroidism in adults: cosponsored by the American Association of Clinical Endocrinologists and the American Thyroid Association. Endocr Pract. 2012;18(6):988–1028.

[26] Demers LM, Spencer CA. Laboratory medicine practice guidelines: laboratory support for the diagnosis and monitoring of thyroid disease. Clin Endocrinol. 2003;58(2):138–140.

[27] Esfandiari NH, Papaleontiou M. Biochemical testing in thyroid disorders. Endocrinol Metab Clin North Am. 2017;46(3):631–648.

[28] Plebani M, Giovanella L. Reflex TSH strategy: the good, the bad and the ugly. Clin Chem Lab Med. 2019;58(1):1–2.

[29] Jonklaas J, Razvi S. Reference intervals in the diagnosis of thyroid dysfunction: treating patients not numbers. Lancet Diabetes Endocrinol. 2019;7(6):473–483.

[30] Biondi B, Cappola AR. Subclinical hypothyroidism in older individuals. Lancet Diabetes Endocrinol. 2022;10(2):129–141.

[31] Lee JH, Kim EY. Resistance to thyroid hormone due to a novel mutation of thyroid hormone receptor beta gene. Ann Pediatr Endocrinol Metab. 2014;19(4):229–231.

[32] Weiss RE, Dumitrescu AM, Refetoff S. Resistance to thyroid hormone and other defects in thyroid

hormone action. MediLib; 2022. Available from: https://www.medilib.ir/uptodate/show/5827.
[33] Wouters H, Slagter SN, Muller Kobold AC, et al. Epidemiology of thyroid disorders in the Lifelines Cohort Study (the Netherlands). PLoS One. 2020; 15(11):e0242795.
[34] Farwell AP. Nonthyroidal illness syndrome. Curr Opin Endocrinol Diabetes Obes. 2013;20(5):478–484.
[35] Gutch M, Kumar S, Gupta KK. Prognostic value of thyroid profile in critical care condition. Indian J Endocrinol Metab. 2018;22(3):387–391.
[36] de Vries EM, Fliers E, Boelen A. The molecular basis of the non-thyroidal illness syndrome. J Endocrinol. 2015;225(3):R67–81.
[37] Kahaly GJ, Bartalena L, Hegedüs L, et al. 2018 European Thyroid Association guideline for the management of Graves' hyperthyroidism. Eur Thyroid J. 2018;7(4):167–186.
[38] Reid JR, Wheeler SF. Hyperthyroidism: diagnosis and treatment. Am Fam Physician. 2005;72(4):623–630.
[39] National Institute for Health and Care Excellence (NICE). Thyroid disease: assessment and management 2019 [cited June 2022]. Available from https://www.nice.org.uk/guidance/ng145.
[40] Henze M, Brown SJ, Hadlow NC, et al. Rationalizing thyroid function testing: which TSH cutoffs are optimal for testing free T4? J Clin Endocrinol Metab. 2017;102(11):4235–4241.
[41] Soldin OP, Soldin SJ. Thyroid hormone testing by tandem mass spectrometry. Clin Biochem. 2011;44(1): 89–94.
[42] Köhrle J, Richards KH. Mass spectrometry-based determination of thyroid hormones and their metabolites in endocrine diagnostics and biomedical research – implications for human serum diagnostics. Exp Clin Endocrinol Diabetes. 2020;128(6–07): 358–374.
[43] Faix JD. Principles and pitfalls of free hormone measurements. Best Pract Res Clin Endocrinol Metab. 2013;27(5):631–645.
[44] Ekins R. Measurement of free hormones in blood. Endocr Rev. 1990;11(1):5–46.
[45] Christofides ND. Free analyte immunoassay. In: Wild D, ed. The immunoassay handbook: theory and applications of ligand binding, ELISA and related techniques. 4th ed. Oxford UK: Elsevier Ltd; 2013. p. 123–138.
[46] Midgley JE. Direct and indirect free thyroxine assay methods: theory and practice. Clin Chem. 2001;47(8): 1353–1363.
[47] Thyroid Disease Manager. Guidelines for diagnosis and management of thyroid disease 2022. [cited 2022 Mar 11]. Available from: https://www.thyroidmanager.org/guidelines/.
[48] Kratzsch J, Baumann NA, Ceriotti F, et al. Global FT4 immunoassay standardization: an expert opinion review. Clin Chem Lab Med. 2021;59(6):1013–1023.
[49] Faix JD, Rosen HN, Velazquez FR. Indirect estimation of thyroid hormone-binding proteins to calculate free thyroxine index: comparison of nonisotopic methods that use labeled thyroxine ("T-uptake"). Clin Chem. 1995;41(1):41–47.
[50] Nusynowitz ML. Free-thyroxine index. JAMA. 1975; 232(10):1050.
[51] Holm SS, Hansen SH, Faber J, et al. Reference methods for the measurement of free thyroid hormones in blood: evaluation of potential reference methods for free thyroxine. Clin Biochem. 2004;37(2):85–93.
[52] Ekins R. Analytic measurements of free thyroxine. Clin Lab Med. 1993;13(3):599–630.
[53] Nelson JC, Tomei RT. Direct determination of free thyroxin in undiluted serum by equilibrium dialysis/radioimmunoassay. Clin Chem. 1988;34(9):1737–1744.
[54] Stockigt JR. Free thyroid hormone measurement. A critical appraisal. Endocrinol Metab Clin North Am. 2001;30(2):265–289.
[55] Nelson JC, Weiss RM. The effect of serum dilution on free thyroxine (T4) concentration in the low T4 syndrome of nonthyroidal illness. J Clin Endocrinol Metab. 1985;61(2):239–246.
[56] Wong TK, Pekary AE, Hoo GS, et al. Comparison of methods for measuring free thyroxin in nonthyroidal illness. Clin Chem. 1992;38(5):720–724.
[57] Soldin SJ, Soukhova N, Janicic N, et al. The measurement of free thyroxine by isotope dilution tandem mass spectrometry. Clin Chim Acta. 2005;358(1–2): 113–118.
[58] Soldin OP, Jang M, Guo T, et al. Pediatric reference intervals for free thyroxine and free triiodothyronine. Thyroid. 2009;19(7):699–702.
[59] Jonklaas J, Kahric-Janicic N, Soldin OP, et al. Correlations of free thyroid hormones measured by tandem mass spectrometry and immunoassay with thyroid-stimulating hormone across 4 patient populations. Clin Chem. 2009;55(7):1380–1388.
[60] Holm SS, Andreasen L, Hansen SH, et al. Influence of adsorption and deproteination on potential free thyroxine reference methods. Clin Chem. 2002;48(1): 108–114.
[61] Grebe SKG. Laboratory testing in thyroid disorders. In Luster M, Duntas LH, Wartofsky L, editors. The thyroid and its diseases: a comprehensive guide for the clinician. 1st ed. Oxford UK: Springer; 2019. p. 129–159.
[62] Favresse J, Burlacu MC, Maiter D, et al. Interferences with thyroid function immunoassays: clinical implications and detection algorithm. Endocr Rev. 2018; 39(5):830–850.
[63] Revet I, Boesten LSM, Linthorst J, et al. Misleading FT4 measurement: assay-dependent antibody interference. Biochem Med. 2016;26(3):436–443.
[64] Migliardi M, Fortunato A, De Renzi G. La misura degli ormoni liberi. Chapter 10. In Dotti C, Fortunato A, editors. Le analisi immunometriche: basi teoriche e applicazioni cliniche. 1st ed. Padova, Italy: Piccin Nuova Libreria; 2014. p. 297–298.
[65] Fritz KS, Wilcox RB, Nelson JC. A direct free thyroxine (T4) immunoassay with the characteristics of a total T4 immunoassay. Clin Chem. 2007;53(5):911–915.
[66] Fritz KS, Wilcox RB, Nelson JC. Quantifying spurious free T4 results attributable to thyroxine-binding proteins in serum dialysates and ultrafiltrates. Clin Chem. 2007;53(5):985–988.

[67] Fillée C, Cumps J, Ketelslegers JM. Comparison of three free T4 (FT4) and free T3 (FT3) immunoassays in healthy subjects and patients with thyroid diseases and severe non-thyroidal illnesses. Clin Lab. 2012;58(7–8):725–736.
[68] Midgley JE, Christofides ND. Point: legitimate and illegitimate tests of free-analyte assay function. Clin Chem. 2009;55(3):439–441.
[69] Gu J, Soldin OP, Soldin SJ. Simultaneous quantification of free triiodothyronine and free thyroxine by isotope dilution tandem mass spectrometry. Clin Biochem. 2007;40(18):1386–1391.
[70] Kuzmanovska S, Miladinova D. Comparison of thyroid-stimulating hormone and free thyroxine immunoassays performed on Immulite 2000 and Maglumi 800 automated analyzers. Open Access Maced J Med Sci. 2020;8(B):168–174.
[71] Padoan A, Cosma C, Plebani M. Evaluation of the analytical performances of six measurands for thyroid functions of Mindray CL-2000i system. J Lab Precis Med. 2018;3.
[72] Yue B, Rockwood AL, Roberts WL. Analysis of free thyroxine in serum by ED-LC-MS/MS. Clin Chem. 2007;53:D122.
[73] Yue B, Rockwood AL, Sandrock T, et al. Free thyroid hormones in serum by direct equilibrium dialysis and online solid-phase extraction–liquid chromatography/tandem mass spectrometry. Clin Chem. 2008; 54(4):642–651.
[74] Köhrle J. The colorful diversity of thyroid hormone metabolites. Eur Thyroid J. 2019;8(3):115–129.
[75] Welsh KJ, Soldin SJ. Diagnosis of endocrine disease: how reliable are free thyroid and total T3 hormone assays? Eur J Endocrinol. 2016;175(6):R255–R263.
[76] Tanoue R, Kume I, Yamamoto Y, et al. Determination of free thyroid hormones in animal serum/plasma using ultrafiltration in combination with ultra-fast liquid chromatography-tandem mass spectrometry. J Chromatogr A. 2018;1539:30–40.
[77] Richards KH, Schanze N, Monk R, et al. A validated LC-MS/MS method for cellular thyroid hormone metabolism: uptake and turnover of mono-iodinated thyroid hormone metabolites by PCCL3 thyrocytes. PLoS One. 2017;12(8):e0183482.
[78] Bowerbank SL, Carlin MG, Dean JR. A direct comparison of liquid chromatography-mass spectrometry with clinical routine testing immunoassay methods for the detection and quantification of thyroid hormones in blood serum. Anal Bioanal Chem. 2019; 411(13):2839–2853.
[79] Annesley TM. Ion suppression in mass spectrometry. Clin Chem. 2003;49(7):1041–1044.
[80] Soukhova N, Soldin OP, Soldin SJ. Isotope dilution tandem mass spectrometric method for T4/T3. Clin Chim Acta. 2004;343(1-2):185–190.
[81] Welsh KJ, Stolze BR, Yu X, et al. Assessment of thyroid function in intensive care unit patients by liquid chromatography tandem mass spectrometry methods. Clin Biochem. 2017;50(6):318–322.
[82] Csako G, Zweig MH, Benson C, et al. On the albumin-dependence of measurements of free thyroxin. II. Patients with non-thyroidal illness. Clin Chem. 1987;33(1):87–92.
[83] Iitaka M, Kawasaki S, Sakurai S, et al. Serum substances that interfere with thyroid hormone assays in patients with chronic renal failure. Clin Endocrinol. 1998;48(6):739–746.
[84] Gant Kanegusuku A, Araque KA, Nguyen H, et al. The effect of specific binding proteins on immunoassay measurements of total and free thyroid hormones and cortisol. Ther Adv Endocrinol. 2021;12: 204201882198924.
[85] Araque KA, Klubo-Gwiezdzinska J, Nieman LK, et al. Assessment of thyroid function tests and harmonization: opinion on thyroid hormone harmonization. Ther Adv Endocrinol Metab. 2019;10:2042018819897049.
[86] Stagnaro-Green A, Abalovich M, Alexander E, et al. Guidelines of the American Thyroid Association for the diagnosis and management of thyroid disease during pregnancy and postpartum. Thyroid. 2011; 21(10):1081–1125.
[87] Anckaert E, Poppe K, Van Uytfanghe K, et al. FT4 immunoassays may display a pattern during pregnancy similar to the equilibrium dialysis ID-LC/tandem MS candidate reference measurement procedure in spite of susceptibility towards binding protein alterations. Clin Chim Acta. 2010;411(17–18): 1348–1353.
[88] Martel J, Després N, Ahnadi CE, et al. Comparative multicentre study of a panel of thyroid tests using different automated immunoassay platforms and specimens at high risk of antibody interference. Clin Chem Lab Med. 2000;38(8):785–793.
[89] d'Herbomez M, Forzy G, Gasser F, et al. Clinical evaluation of nine free thyroxine assays: persistent problems in particular populations. Clin Chem Lab Med. 2003;41(7):942–947.
[90] Thienpont LM. A major step forward in the routine measurement of serum free thyroid hormones. Clin Chem. 2008;54(4):625–626.
[91] Giovannini S, Zucchelli GC, Iervasi G, et al. Multicentre comparison of free thyroid hormones immunoassays: the Immunocheck study. Clin Chem Lab Med. 2011;49(10):1669–1676.
[92] Kristensen GB, Rustad P, Berg JP, et al. Analytical bias exceeding desirable quality goal in 4 out of 5 common immunoassays: results of a native single serum sample external quality assessment program for cobalamin, folate, ferritin, thyroid-stimulating hormone, and free T4 analyses. Clin Chem. 2016;62(9): 1255–1263.
[93] Tate JR, Yen T, Jones GR. Transference and validation of reference intervals. Clin Chem. 2015;61(8): 1012–1015.
[94] Berg J. The UK Pathology Harmony initiative; the foundation of a global model. Clin Chim Acta. 2014; 432:22–26.
[95] Thienpont LM, Van Uytfanghe K, Van Houcke S, et al. A progress report of the IFCC Committee for Standardization of Thyroid Function Tests. Eur Thyroid J. 2014;3(2):109–116.
[96] Thienpont LM, Van Uytfanghe K, De Grande LAC, et al. Harmonization of serum thyroid-stimulating

hormone measurements paves the way for the adoption of a more uniform reference interval. Clin Chem. 2017;63(7):1248–1260.
[97] Vesper HW, Van Uytfanghe K, Hishinuma A, et al. Implementing reference systems for thyroid function tests – a collaborative effort. Clin Chim Acta. 2021; 519:183–186.
[98] Thienpont LM, Van Uytfanghe K, Van Houcke S. Standardization activities in the field of thyroid function tests: a status report. Clin Chem Lab Med. 2010; 48(11):1577–1583.
[99] International Federation of Clinical Chemistry and Laboratory Medicine (IFCC). Standardization of thyroid function tests (C-STFT) 2022. [cited 2022 Mar 11]. Available from: https://www.ifcc.org/ifcc-scientific-division/sd-committees/c-stft/.
[100] ISO. In vitro diagnostic medical devices — requirements for establishing metrological traceability of values assigned to calibrators, trueness control materials and human samples. ISO 175112020. Geneva, Switzerland: International Organization for Standardization; 2020.
[101] Van Houcke SK, Van Uytfanghe K, Shimizu E, et al. IFCC international conventional reference procedure for the measurement of free thyroxine in serum: International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) Working Group for Standardization of Thyroid Function Tests (WG-STFT)(1). Clin Chem Lab Med. 2011;49(8):1275–1281.
[102] CLSI. Measurement of free thyroid hormones. Approved guideline. CLSI document C45-A. Wayne, PA: Clinical and Laboratory Standards Institute; 2004.
[103] IFCC. Reference intervals and decision limits (C-RIDL). Geneva, Switzerland: International Federation of Clinical Chemistry and Laboratory Medicine; 2022. Available from: https://www.ifcc.org/ifcc-scientific-division/sd-committees/c-ridl/
[104] IFCC. Summary of activities to collect information about concerns and potential risks associated with changes in reference intervals for free thyroxine and TSH and about communication and interactions among relevant stakeholders (C-STFT). Geneva, Switzerland: International Federation of Clinical Chemistry and Laboratory Medicine; 2022. Available from: https://ifcc-cstft.org/research/summary-of-activities-to-collect-information-about-concerns-and-potential-risks-associated.
[105] Midgley JE. Global FT4 immunoassay standardization. Response to: Kratzsch J et al. Global FT4 immunoassay standardization: an expert opinion review. Clin Chem Lab Med. 2021;59(6):e223–e224.
[106] Giovanella L. Free-thyroxine standardization: waiting for godot while well serving our patients today. Clin Chem Lab Med. 2021;59(6):e225–e226.
[107] Jones GRD, Haeckel R, Loh TP, et al. Indirect methods for reference interval determination – review and recommendations. Clin Chem Lab Med. 2018;57(1): 20–29.
[108] CLSI. Defining, establishing, and verifying reference intervals in the clinical laboratory. Approved Guideline – Third Edition. CLSI document EP28-A3C. Wayne, PA: Clinical and Laboratory Standards Institute; 2010.
[109] Zou Y, Wang D, Cheng X, et al. Reference intervals for thyroid-associated hormones and the prevalence of thyroid diseases in the Chinese population. Ann Lab Med. 2021;41(1):77–85.
[110] Ceriotti F, Henny J. "Are my laboratory results normal?" Considerations to be made concerning reference intervals and decision limits. Ejifcc. 2008;19(2): 106–114.
[111] Ceriotti F, Hinzmann R, Panteghini M. Reference intervals: the way forward. Ann Clin Biochem. 2009; 46(Pt 1):8–17.
[112] Sikaris KA. Physiology and its importance for reference intervals. Clin Biochem Rev. 2014;35(1):3–14.
[113] Kouri T, Kairisto V, Virtanen A, et al. Reference intervals developed from data for hospitalized patients: computerized method based on combination of laboratory and diagnostic data. Clin Chem. 1994; 40(12):2209–2215.
[114] Bhattacharya CG. A simple method of resolution of a distribution into gaussian components. Biometrics. 1967;23(1):115–135.
[115] Jones G, Horowitz G, Katayev A, et al. Reference intervals data mining: getting the right paper. Am J Clin Pathol. 2015;144(3):526–527.
[116] Yang J, Li Y, Liu Q, et al. Brief introduction of medical database and data mining technology in big data era. J Evid Based Med. 2020;13(1):57–69.
[117] Barth JH, Luvai A, Jassam N, et al. Comparison of method-related reference intervals for thyroid hormones: studies from a prospective reference population and a literature review. Ann Clin Biochem. 2018; 55(1):107–112.
[118] Quinn FA, Tam MC, Wong PT, et al. Thyroid autoimmunity and thyroid hormone reference intervals in apparently healthy Chinese adults. Clin Chim Acta. 2009;405(1–2):156–159.
[119] Schalin-Jäntti C, Tanner P, Välimäki MJ, et al. Serum TSH reference interval in healthy Finnish adults using the Abbott Architect 2000i analyzer. Scand J Clin Lab Invest. 2011;71(4):344–349.
[120] Milinković N, Ignjatović S, Žarković M, et al. Indirect estimation of age-related reference limits of thyroid parameters: a cross-sectional study of outpatients' results. Scand J Clin Lab Invest. 2014;74(5):378–384.
[121] De Grande LAC, Van Uytfanghe K, Reynders D, et al. Standardization of free thyroxine measurements allows the adoption of a more uniform reference interval. Clin Chem. 2017;63(10):1642–1652.
[122] Ehrenkranz J, Bach PR, Snow GL, et al. Circadian and circannual rhythms in thyroid hormones: determining the TSH and free T4 reference intervals based upon time of day, age, and sex. Thyroid. 2015;25(8): 954–961.
[123] Hickman PE, Koerbin G, Simpson A, et al. Using a thyroid disease-free population to define the reference interval for TSH and free T4 on the Abbott Architect analyser. Clin Endocrinol. 2017;86(1): 108–112.
[124] Barhanovic NG, Antunovic T, Kavaric S, et al. Age and assay related changes of laboratory thyroid

function tests in the reference female population. J Med Biochem. 2019;38(1):22–32.
[125] Plouvier E, Alliot L, Bigorie B, et al. [Necessity to define thyroid reference values for better clinical interpretation]. Ann Biol Clin. 2011;69(1):77–83.
[126] Baloch Z, Carayon P, Conte-Devolx B, et al. Laboratory medicine practice guidelines. Laboratory support for the diagnosis and monitoring of thyroid disease. Thyroid. 2003;13(1):3–126.
[127] Wang Y, Zhang YX, Zhou YL, et al. Establishment of reference intervals for serum thyroid-stimulating hormone, free and total thyroxine, and free and total triiodothyronine for the Beckman Coulter DxI-800 analyzers by indirect method using data obtained from Chinese population in Zhejiang province. J Clin Lab Anal. 2017;31(4):e22069.
[128] Li ZZ, Yu BZ, Wang JL, et al. Reference intervals for thyroid-stimulating hormone and thyroid hormones using the access TSH 3rd IS method in China. J Clin Lab Anal. 2020;34(5):e23197.
[129] Wang D, Yu S, Cheng X, et al. Nationwide Chinese study for establishing reference intervals for thyroid hormones and related tests. Clin Chim Acta. 2019;496:62–67.
[130] Kratzsch J, Fiedler GM, Leichtle A, et al. New reference intervals for thyrotropin and thyroid hormones based on National Academy of Clinical Biochemistry criteria and regular ultrasonography of the thyroid. Clin Chem. 2005;51(8):1480–1486.
[131] Friis-Hansen L, Hilsted L. Reference intervals for thyreotropin and thyroid hormones for healthy adults based on the NOBIDA material and determined using a Modular E170. Clin Chem Lab Med. 2008;46(9):1305–1312.
[132] Takeda K, Mishiba M, Sugiura H, et al. Evaluated reference intervals for serum free thyroxine and thyrotropin using the conventional outliner rejection test without regard to presence of thyroid antibodies and prevalence of thyroid dysfunction in Japanese subjects. Endocr J. 2009;56(9):1059–1066.
[133] Chan AO, Iu YP, Shek CC. The reference interval of thyroid-stimulating hormone in Hong Kong Chinese. J Clin Pathol. 2011;64(5):433–436.
[134] Yoshihara A, Noh JY, Ohye H, et al. Reference limits for serum thyrotropin in a Japanese population. Endocr J. 2011;58(7):585–588.
[135] Amouzegar A, Delshad H, Mehran L, et al. Reference limit of thyrotropin (TSH) and free thyroxine (FT4) in thyroperoxidase positive and negative subjects: a population based study. J Endocrinol Invest. 2013;36(11):950–954.
[136] Marwaha RK, Tandon N, Ganie MA, et al. Reference range of thyroid function (FT3, FT4 and TSH) among Indian adults. Clin Biochem. 2013;46(4–5):341–345.
[137] Fontes R, Coeli CR, Aguiar F, et al. Reference interval of thyroid stimulating hormone and free thyroxine in a reference population over 60 years old and in very old subjects (over 80 years): comparison to young subjects. Thyroid Res. 2013;6(1):13.
[138] Sriphrapradang C, Pavarangkoon S, Jongjaroenprasert W, et al. Reference ranges of serum TSH, FT4 and thyroid autoantibodies in the Thai population: the national health examination survey. Clin Endocrinol (Oxf). 2014;80(5):751–756.
[139] Mirjanic-Azaric B, Avram S, Stojakovic-Jelisavac T, et al. Direct estimation of reference intervals for thyroid parameters in the Republic of Srpska. J Med Biochem. 2017;36(2):137–144.
[140] Park SY, Kim HI, Oh HK, et al. Age- and gender-specific reference intervals of TSH and free T4 in an iodine-replete area: data from Korean National Health and Nutrition Examination Survey IV (2013-2015). PLoS One. 2018;13(2):e0190738.
[141] Reix N, Massart C, d'Herbomez M, et al. Thyroid-stimulating hormone and free thyroxine on the ADVIA centaur immunoassay system: a multicenter assessment of analytical performance. Clin Biochem. 2013;46(13–14):1305–1308.
[142] Wang P, Gao YJ, Cheng J, et al. Serum thyroid hormone reference intervals in the apparently healthy individuals of Zhengzhou area of China. Genet Mol Res. 2014;13(3):7275–7281.
[143] Cai J, Fang Y, Jing D, et al. Reference intervals of thyroid hormones in a previously iodine-deficient but presently more than adequate area of Western China: a population-based survey. Endocr J. 2016;63(4):381–388.
[144] Hoermann R, Larisch R, Dietrich JW, et al. Derivation of a multivariate reference range for pituitary thyrotropin and thyroid hormones: diagnostic efficiency compared with conventional single-reference method. Eur J Endocrinol. 2016;174(6):735–743.
[145] Barth JH, Spencer JD, Goodall SR, et al. Reference intervals for thyroid hormones on Advia Centaur derived from three reference populations and a review of the literature. Ann Clin Biochem. 2016;53(Pt 3):385–389.
[146] Jones GR, Koetsier SD. RCPAQAP first combined measurement and reference interval survey. Clin Biochem Rev. 2014;35(4):243–250.
[147] Lee GR, Griffin A, Halton K, et al. Generating method-specific reference ranges – a harmonious outcome? Pract Lab Med. 2017;9:1–11.
[148] Kratzsch J, Schubert G, Pulzer F, et al. Reference intervals for TSH and thyroid hormones are mainly affected by age, body mass index and number of blood leucocytes, but hardly by gender and thyroid autoantibodies during the first decades of life. Clin Biochem. 2008;41(13):1091–1098.
[149] Tahmasebi H, Trajcevski K, Higgins V, et al. Influence of ethnicity on population reference values for biochemical markers. Crit Rev Clin Lab Sci. 2018;55(5):359–375.
[150] Bohn MK, Higgins V, Asgari S, et al. Paediatric reference intervals for 17 Roche cobas 8000 e602 immunoassays in the CALIPER cohort of healthy children and adolescents. Clin Chem Lab Med. 2019;57(12):1968–1979.
[151] Strich D, Karavani G, Levin S, et al. Normal limits for serum thyrotropin vary greatly depending on method. Clin Endocrinol. 2016;85(1):110–115.
[152] Yeap BB, Manning L, Chubb SA, et al. Reference ranges for thyroid-stimulating hormone and free thyroxine in older men: results from the Health In Men

study. J Gerontol A Biol Sci Med Sci. 2017;72(3):444–449.
[153] Boelaert K, Torlinska B, Holder RL, et al. Older subjects with hyperthyroidism present with a paucity of symptoms and signs: a large cross-sectional study. J Clin Endocrinol Metab. 2010;95(6):2715–2726.
[154] Mitrou P, Raptis SA, Dimitriadis G. Thyroid disease in older people. Maturitas. 2011;70(1):5–9.
[155] Xiong J, Liu S, Hu K, et al. Study of reference intervals for free triiodothyronine, free thyroxine, and thyroid-stimulating hormone in an elderly Chinese Han population. PLoS One. 2020;15(9):e0239579.
[156] Gussekloo J, van Exel E, de Craen AJ, et al. Thyroid status, disability and cognitive function, and survival in old age. JAMA. 2004;292(21):2591–2599.
[157] Yeap BB, Alfonso H, Hankey GJ, et al. Higher free thyroxine levels are associated with all-cause mortality in euthyroid older men: the Health In Men study. Eur J Endocrinol. 2013;169(4):401–408.
[158] Correia A, Nascimento MLF, Teixeira L, et al. Free thyroxine but not TSH levels are associated with decline in functional status in a cohort of geriatric outpatients. Eur Geriatr Med. 2022;13(1):147–154.
[159] Wang D, Yu S, Ma C, et al. Reference intervals for thyroid-stimulating hormone, free thyroxine, and free triiodothyronine in elderly Chinese persons. Clin Chem Lab Med. 2019;57(7):1044–1052.
[160] Yazbeck CF, Sullivan SD. Thyroid disorders during pregnancy. Med Clin North Am. 2012;96(2):235–256.
[161] De Groot L, Abalovich M, Alexander EK, et al. Management of thyroid dysfunction during pregnancy and postpartum: an Endocrine Society clinical practice guideline. J Clin Endocrinol Metab. 2012;97(8):2543–2565.
[162] Ekinci EI, Lu ZX, Sikaris K, et al. Longitudinal assessment of thyroid function in pregnancy. Ann Clin Biochem. 2013;50(Pt 6):595–602.
[163] Springer D, Jiskra J, Limanova Z, et al. Thyroid in pregnancy: from physiology to screening. Crit Rev Clin Lab Sci. 2017;54(2):102–116.
[164] Glinoer D. The regulation of thyroid function in pregnancy: pathways of endocrine adaptation from physiology to pathology. Endocr Rev. 1997;18(3):404–433.
[165] Alexander EK, Pearce EN, Brent GA, et al. 2017 Guidelines of the American Thyroid Association for the diagnosis and management of thyroid disease during pregnancy and the postpartum. Thyroid. 2017;27(3):315–389.
[166] Ho CKM, Tan ETH, Ng MJ, et al. Gestational age-specific reference intervals for serum thyroid hormone levels in a multi-ethnic population. Clin Chem Lab Med. 2017;55(11):1777–1788.
[167] Veltri F, Belhomme J, Kleynen P, et al. Maternal thyroid parameters in pregnant women with different ethnic backgrounds: do ethnicity-specific reference ranges improve the diagnosis of subclinical hypothyroidism? Clin Endocrinol. 2017;86(6):830–836.
[168] Korevaar TI, Medici M, de Rijke YB, et al. Ethnic differences in maternal thyroid parameters during pregnancy: the Generation R study. J Clin Endocrinol Metab. 2013;98(9):3678–3686.
[169] Mosso L, Martínez A, Rojas MP, et al. Early pregnancy thyroid hormone reference ranges in Chilean women: the influence of body mass index. Clin Endocrinol. 2016;85(6):942–948.
[170] Collares FM, Korevaar TIM, Hofman A, et al. Maternal thyroid function, prepregnancy obesity and gestational weight gain-the Generation R study: a prospective cohort study. Clin Endocrinol. 2017;87(6):799–806.
[171] Wiersinga WM. Smoking and thyroid. Clin Endocrinol. 2013;79(2):145–151.
[172] Joosen AM, van der Linden IJ, de Jong-Aarts N, et al. TSH and fT4 during pregnancy: an observational study and a review of the literature. Clin Chem Lab Med. 2016;54(7):1239–1246.
[173] Okosieme OE, Agrawal M, Usman D, et al. Method-dependent variation in TSH and FT4 reference intervals in pregnancy: a systematic review. Ann Clin Biochem. 2021;58(5):537–546.
[174] Akarsu S, Akbiyik F, Karaismailoglu E, et al. Gestation specific reference intervals for thyroid function tests in pregnancy. Clin Chem Lab Med. 2016;54(8):1377–1383.
[175] Fan JX, Yang S, Qian W, et al. Comparison of the reference intervals used for the evaluation of maternal thyroid function during pregnancy using sequential and nonsequential methods. Chin Med J (Engl). 2016;129(7):785–791.
[176] Hernández JM, Soldevila B, Velasco I, et al. Reference intervals of thyroid function tests assessed by immunoassay and mass spectrometry in healthy pregnant women living in Catalonia. JCM. 2021;10(11):2444.
[177] Liu J, Yu X, Xia M, et al. Development of gestation-specific reference intervals for thyroid hormones in normal pregnant Northeast Chinese women: what is the rational division of gestation stages for establishing reference intervals for pregnancy women? Clin Biochem. 2017;50(6):309–317.
[178] Ollero MD, Toni M, Pineda JJ, et al. Thyroid function reference values in healthy iodine-sufficient pregnant women and influence of thyroid nodules on thyrotropin and free thyroxine values. Thyroid. 2019;29(3):421–429.
[179] Quinn FA, Reyes-Mendez MA, Nicholson L, et al. Thyroid function and thyroid autoimmunity in apparently healthy pregnant and non-pregnant Mexican women. Clin Chem Lab Med. 2014;52(9):1305–1311.
[180] Shen FX, Xie ZW, Lu SM, et al. Gestational thyroid reference intervals in antibody-negative Chinese women. Clin Biochem. 2014;47(7–8):673–675.
[181] Yang X, Meng Y, Zhang Y, et al. Thyroid function reference ranges during pregnancy in a large Chinese population and comparison with current guidelines. Chin Med J. 2019;132(5):505–511.
[182] Yuen LY, Chan MHM, Sahota DS, et al. Development of gestational age-specific thyroid function test reference intervals in four analytic platforms through multilevel modeling. Thyroid. 2020;30(4):598–608.
[183] Bunch DR, Firmender K, Harb R, et al. First- and second-trimester reference intervals for thyroid

function testing in a US population. Am J Clin Pathol. 2021;155(6):776–780.
[184] Donovan LE, Metcalfe A, Chin A, et al. A practical approach for the verification and determination of site- and trimester-specific reference intervals for thyroid function tests in pregnancy. Thyroid. 2019;29(3):412–420.
[185] Duan Y, Peng L, Cui Y, et al. Reference intervals for thyroid function and the negative correlation between FT4 and HbA1c in pregnant women of West China. Clin Lab. 2015;61(7):777–783.
[186] Han L, Zheng W, Zhai Y, et al. Reference intervals of trimester-specific thyroid stimulating hormone and free thyroxine in Chinese women established by experimental and statistical methods. J Clin Lab Anal. 2018;32(4):e22344.
[187] Huang C, Wu Y, Chen L, et al. Establishment of assay method- and trimester-specific reference intervals for thyroid hormones during pregnancy in Chengdu. J Clin Lab Anal. 2021;35(5):e23763.
[188] Khalil AB, Salih BT, Chinengo O, et al. Trimester specific reference ranges for serum TSH and free T4 among United Arab Emirates pregnant women. Pract Lab Med. 2018;12:e00098.
[189] Kim HJ, Cho YY, Kim SW, et al. Reference intervals of thyroid hormones during pregnancy in Korea, an iodine-replete area. Korean J Intern Med. 2018;33(3):552–560.
[190] Moon HW, Chung HJ, Park CM, et al. Establishment of trimester-specific reference intervals for thyroid hormones in Korean pregnant women. Ann Lab Med. 2015;35(2):198–204.
[191] Sekhri T, Juhi JA, Wilfred R, et al. Trimester specific reference intervals for thyroid function tests in normal Indian pregnant women. Indian J Endocrinol Metab. 2016;20(1):101–107.
[192] Sun R, Xia J. The reference intervals of thyroid hormones for pregnant women in Zhejiang province. Lab Med. 2017;49(1):5–10.
[193] Zhang J, Li W, Chen QB, et al. Establishment of trimester-specific thyroid stimulating hormone and free thyroxine reference interval in pregnant Chinese women using the Beckman Coulter UniCel™ DxI 600. Clin Chem Lab Med. 2015;53(9):1409–1414.
[194] Zhang X, Yao B, Li C, et al. Reference intervals of thyroid function during pregnancy: self-sequential longitudinal study versus cross-sectional study. Thyroid. 2016;26(12):1786–1793.
[195] Zhou H, Ma ZF, Lu Y, et al. Assessment of iodine status among pregnant women and neonates using neonatal thyrotropin (TSH) in mainland China after the introduction of new revised universal salt iodisation (USI) in 2012: a re-emergence of iodine deficiency? Int J Endocrinol. 2019;2019:3618169.
[196] Zhang D, Cai K, Wang G, et al. Trimester-specific reference ranges for thyroid hormones in pregnant women. Medicine. 2019;98(4):e14245.
[197] Forhead AJ, Fowden AL. Thyroid hormones in fetal growth and prepartum maturation. J Endocrinol. 2014;221(3):R87–R103.
[198] Derakhshan A, Shu H, Broeren MAC, et al. Reference ranges and determinants of thyroid function during early pregnancy: the SELMA study. J Clin Endocrinol Metab. 2018;103(9):3548–3556.
[199] Andersen SL, Christensen PA, Knøsgaard L, et al. Classification of thyroid dysfunction in pregnant women differs by analytical method and type of thyroid function test. J Clin Endocrinol Metab. 2020;105:dgaa567.
[200] Bliddal S, Feldt-Rasmussen U, Boas M, et al. Gestational age-specific reference ranges from different laboratories misclassify pregnant women's thyroid status: comparison of two longitudinal prospective cohort studies. Eur J Endocrinol. 2014;170(2):329–339.
[201] Lee RH, Spencer CA, Mestman JH, et al. Free T4 immunoassays are flawed during pregnancy. Am J Obstet Gynecol. 2009;200(3):260.e261–266.
[202] Geno KA, Reed MS, Cervinski MA, et al. Evaluation of thyroid function in pregnant women using automated immunoassays. Clin Chem. 2021;67(5):772–780.
[203] Feldt-Rasmussen U, Bliddal S, Rasmussen AK, et al. Challenges in interpretation of thyroid function tests in pregnant women with autoimmune thyroid disease. J Thyroid Res. 2011;2011:1–7.
[204] Tarım Ö. Thyroid hormones and growth in health and disease. J Clin Res Pediatr Endocrinol. 2011;3(2):51–55.
[205] Leung AKC, Leung AAC. Evaluation and management of the child with hypothyroidism. World J Pediatr. 2019;15(2):124–134.
[206] Srinivasan S, Misra M. Hyperthyroidism in children. Pediatr Rev. 2015;36(6):239–248.
[207] Chaudhari M, Slaughter JL. Thyroid function in the neonatal intensive care unit. Clin Perinatol. 2018;45(1):19–30.
[208] Lazarus J, Brown RS, Daumerie C, et al. 2014 European Thyroid Association guidelines for the management of subclinical hypothyroidism in pregnancy and in children. Eur Thyroid J. 2014;3(2):76–94.
[209] Adeli K, Higgins V, Trajcevski K, et al. The Canadian laboratory initiative on pediatric reference intervals: a CALIPER white paper. Crit Rev Clin Lab Sci. 2017;54(6):358–413.
[210] Colantonio DA, Kyriakopoulou L, Chan MK, et al. Closing the gaps in pediatric laboratory reference intervals: a CALIPER database of 40 biochemical markers in a healthy and multiethnic population of children. Clin Chem. 2012;58(5):854–868.
[211] Bailey D, Colantonio D, Kyriakopoulou L, et al. Marked biological variance in endocrine and biochemical markers in childhood: establishment of pediatric reference intervals using healthy community children from the CALIPER cohort. Clin Chem. 2013;59(9):1393–1405.
[212] Karbasy K, Lin DC, Stoianov A, et al. Pediatric reference value distributions and covariate-stratified reference intervals for 29 endocrine and special chemistry biomarkers on the beckman coulter immunoassay systems: a CALIPER study of healthy community children. Clin Chem Lab Med. 2016;54(4):643–657.
[213] Higgins V, Fung AWS, Chan MK, et al. Pediatric reference intervals for 29 Ortho VITROS 5600

immunoassays using the CALIPER cohort of healthy children and adolescents. Clin Chem Lab Med. 2018; 56(2):327–340.
[214] Bohn MK, Horn P, League D, et al. Pediatric reference intervals for endocrine markers and fertility hormones in healthy children and adolescents on the Siemens Healthineers Atellica immunoassay system. Clin Chem Lab Med. 2021;59(8):1421–1430.
[215] Aldrimer M, Ridefelt P, Rödöö P, et al. Reference intervals on the Abbott Architect for serum thyroid hormones, lipids and prolactin in healthy children in a population-based study. Scand J Clin Lab Invest. 2012;72(4):326–332.
[216] Önsesveren I, Barjaktarovic M, Chaker L, et al. Childhood thyroid function reference ranges and determinants: a literature overview and a prospective cohort study. Thyroid. 2017;27(11):1360–1369.
[217] Argente Del Castillo P, Pastor García MI, Morell-Garcia D, et al. Thyroid panel reference intervals in healthy children and adolescents: a Spanish cohort. Clin Biochem. 2021;91:39–44.
[218] Bokulić A, Zec I, Marijančević D, et al. Establishing paediatric reference intervals for thyroid function tests in Croatian population on the Abbott Architect i2000. Biochem Med. 2021;31(3):030702.
[219] Radicioni AF, Tahani N, Spaziani M, et al. Reference ranges for thyroid hormones in normal Italian children and adolescents and overweight adolescents. J Endocrinol Invest. 2013;36(5):326–330.
[220] Romero-Villarreal JB, Palacios-Saucedo GC, Ocaña-Hernández LA, et al. [Reference intervals for total triiodothyronine (TT3), free thyroxine (FT4), and thyrotropin (TSH) by chemiluminescence immunoassay in children younger than six years old from northeastern Mexico]. Gac Med Mex. 2014;150(Suppl 2): 248–254.
[221] Gunapalasingham G, Frithioff-Bøjsøe C, Lund MAV, et al. Reference values for fasting serum concentrations of thyroid-stimulating hormone and thyroid hormones in Healthy Danish/North-European white children and adolescents. Scand J Clin Lab Invest. 2019;79(1–2):129–135.
[222] Iwaku K, Noh JY, Minagawa A, et al. Determination of pediatric reference levels of FT3, FT4 and TSH measured with ECLusys kits. Endocr J. 2013;60(6): 799–804.
[223] Loh TP, Sethi SK, Metz MP. Paediatric reference interval and biological variation trends of thyrotropin (TSH) and free thyroxine (T4) in an Asian population. J Clin Pathol. 2015;68(8):642–647.
[224] Strich D, Edri S, Gillis D. Current normal values for TSH and FT3 in children are too low: evidence from over 11,000 samples. J Pediatr Endocrinol Metab. 2012;25(3–4):245–248.
[225] Kapelari K, Kirchlechner C, Högler W, et al. Pediatric reference intervals for thyroid hormone levels from birth to adulthood: a retrospective study. BMC Endocr Disord. 2008;8(1):15.
[226] Soldin SJ, Cheng LL, Lam LY, et al. Comparison of FT4 with log TSH on the Abbott Architect ci8200: pediatric reference intervals for free thyroxine and thyroid-stimulating hormone. Clin Chim Acta. 2010; 411(3-4):250–252.
[227] Verburg FA, Kirchgässner C, Hebestreit H, et al. Reference ranges for analytes of thyroid function in children. Horm Metab Res. 2011;43(6):422–426.
[228] Oladipo O, Nenninger DA, Parvin CA, et al. Intraindividual variability of thyroid function tests in a pediatric population. Clin Chim Acta. 2010; 411(15–16):1143–1145.
[229] Jayasuriya MS, Choy KW, Chin LK, et al. Reference intervals for neonatal thyroid function tests in the first 7 days of life. J Pediatr Endocrinol Metab. 2018; 31(10):1113–1116.
[230] Naafs JC, Heinen CA, Zwaveling-Soonawala N, et al. Age-specific reference intervals for plasma free thyroxine and thyrotropin in term neonates during the first two weeks of life. Thyroid. 2020;30(8): 1106–1111.
[231] Campbell PJ, Brown SJ, Kendrew P, et al. Changes in thyroid function across adolescence: a longitudinal study. J Clin Endocrinol Metab. 2020;105:dgz331.
[232] Taylor PN, Razvi S, Pearce SH, et al. Clinical review: a review of the clinical consequences of variation in thyroid function within the reference range. J Clin Endocrinol Metab. 2013;98(9):3562–3571.
[233] Fisher DA, Grueters A. Disorders of the thyroid in the newborn and infant. Chapter 6. In Sperling MA, editor. Pediatric endocrinology. 3rd ed. Philadelphia, PA: Saunders Elsevier; 2008. p. 198–226.
[234] Aktas S. Evaluation of thyrotropin and thyroxine levels in the first month of life. Cyprus J Med Sci. 2019; 4(2):99–102.
[235] Mutlu M, Karagüzel G, Alı yazicioğlu Y, et al. Reference intervals for thyrotropin and thyroid hormones and ultrasonographic thyroid volume during the neonatal period. J Matern Fetal Neonatal Med. 2012;25(2):120–124.
[236] Omuse G, Kassim A, Kiigu F, et al. Reference intervals for thyroid stimulating hormone and free thyroxine derived from neonates undergoing routine screening for congenital hypothyroidism at a university teaching hospital in Nairobi, Kenya: a cross sectional study. BMC Endocr Disord. 2016;16(1):23.
[237] Wong JSL, Selveindran NM, Mohamed RZ, et al. Reference intervals for thyroid-stimulating hormone (TSH) and free thyroxine (FT4) in infants' day 14-30 of life and a comparison with other studies. J Pediatr Endocrinol Metab. 2020;33(9):1125–1132.
[238] Lin X, Zheng LJ, Li HB, et al. Reference intervals for preterm thyroid function during the fifth to seventh day of life. Clin Biochem. 2021;95:54–59.
[239] Karavani G, Strich D, Edri S, et al. Increases in thyrotropin within the near-normal range are associated with increased triiodothyronine but not increased thyroxine in the pediatric age group. J Clin Endocrinol Metab. 2014;99(8):E1471–E1475.
[240] Fisher DA, Grueters A. Thyroid disorders in childhood and adolescence. Chapter 7. In Sperling MA, ed. Pediatric endocrinology. 3rd ed. Philadelphia, PA: Saunders Elsevier; 2008. p. 227–253.
[241] Galior K, Stan M, Baumann N. Higher FT4 results in levothyroxine-treated patients with normal TSH

compared to patients without thyroid disease. J Endocr Soc. 2019;3(Suppl_1):MON–623.
[242] Lu ZX, Sikaris KA, Yen T, et al. Should there be separate free thyroxine reference limits for thyroxine-treated patients? Clin Biochem Rev. 2016;37:S40.
[243] Wheeler E, Choy KW, Chin LK, et al. Routine free thyroxine reference intervals are suboptimal for monitoring children on thyroxine replacement therapy and target intervals need to be assay-specific. Sci Rep. 2019;9(1):19080.
[244] Fraser CG. Reference change values. Clin Chem Lab Med. 2011;50(5):807–812.
[245] Fraser CG. Reference change values: the way forward in monitoring. Ann Clin Biochem. 2009;46(Pt 3): 264–265.
[246] Andersen S, Bruun NH, Pedersen KM, et al. Biologic variation is important for interpretation of thyroid function tests. Thyroid. 2003;13(11):1069–1078.
[247] Andersen S, Pedersen KM, Bruun NH, et al. Narrow individual variations in serum T(4) and T(3) in normal subjects: a clue to the understanding of subclinical thyroid disease. J Clin Endocrinol Metab. 2002;87(3): 1068–1072.
[248] Ankrah-Tetteh T, Wijeratne S, Swaminathan R. Intraindividual variation in serum thyroid hormones, parathyroid hormone and insulin-like growth factor-1. Ann Clin Biochem. 2008;45(2):167–169.
[249] Erden G, Barazi AO, Tezcan G, et al. Biological variation and reference change values of TSH, free T3, and free T4 levels in serum of healthy Turkish individuals. Turk J Med Sci. 2008;38:153–158.
[250] Bottani M, Aarsand AK, Banfi G, et al. European biological variation study (EuBIVAS): within- and between-subject biological variation estimates for serum thyroid biomarkers based on weekly samplings from 91 healthy participants. Clin Chem Lab Med. 2022;60(4):523–532.
[251] Mairesse A, Wauthier L, Courcelles L, et al. Biological variation and analytical goals of four thyroid function biomarkers in healthy European volunteers. Clin Endocrinol (Oxf). 2021;94(5):845–850.
[252] Zhang Y, He DH, Jiang SN, et al. Biological variation of thyroid function biomarkers over 24 hours. Clin Chim Acta. 2021;523:519–524.
[253] Ghazal K, Brabant S, Prie D, et al. Hormone immunoassay interference: a 2021 update. Ann Lab Med. 2022;42(1):3–23.
[254] Hattori N, Ishihara T, Shimatsu A. Variability in the detection of macro TSH in different immunoassay systems. Eur J Endocrinol. 2016;174(1):9–15.
[255] Ismail AA, Walker PL, Barth JH, et al. Wrong biochemistry results: two case reports and observational study in 5310 patients on potentially misleading thyroid-stimulating hormone and gonadotropin immunoassay results. Clin Chem. 2002;48(11): 2023–2029.
[256] Drees JC, Stone JA, Reamer CR, et al. Falsely undetectable TSH in a cohort of South Asian euthyroid patients. J Clin Endocrinol Metab. 2014;99(4): 1171–1179.
[257] Estrada JM, Soldin D, Buckey TM, et al. Thyrotropin isoforms: implications for thyrotropin analysis and clinical practice. Thyroid. 2014;24(3):411–423.
[258] Sapin R, Schlienger JL, Grunenberger F, et al. In vitro and in vivo effects of increased concentrations of free fatty acids on free thyroxin measurements as determined by five assays. Clin Chem. 1990;36(4): 611–613.
[259] Toldy E, Locsei Z, Szabolcs I, et al. Protein interference in thyroid assays: an in vitro study with in vivo consequences. Clin Chim Acta. 2005;352(1–2):93–104.
[260] DeCosimo DR, Fang SL, Braverman LE. Prevalence of familial dysalbuminemic hyperthyroxinemia in Hispanics. Ann Intern Med. 1987;107(5):780–781.
[261] Mimoto MS, Refetoff S. Clinical recognition and evaluation of patients with inherited serum thyroid hormone-binding protein mutations. J Endocrinol Invest. 2020;43(1):31–41.
[262] Koulouri O, Moran C, Halsall D, et al. Pitfalls in the measurement and interpretation of thyroid function tests. Best Pract Res Clin Endocrinol Metab. 2013; 27(6):745–762.
[263] Jaume JC, Mendel CM, Frost PH, et al. Extremely low doses of heparin release lipase activity into the plasma and can thereby cause artifactual elevations in the serum-free thyroxine concentration as measured by equilibrium dialysis. Thyroid. 1996;6(2): 79–83.
[264] Külz M, Fellner S, Rocktäschel J, et al. Dubiously increased FT4 and FT3 levels in clinically euthyroid patients: clinical finding or analytical pitfall? Clin Chem Lab Med. 2022;60(6):877–885.
[265] Ghosh S, Howlett M, Boag D, et al. Interference in free thyroxine immunoassay. Eur J Intern Med. 2008; 19(3):221–222.
[266] Monchamp T, Chopra IJ, Wah DT, et al. Falsely elevated thyroid hormone levels due to anti-sheep antibody interference in an automated electrochemiluminescent immunoassay. Thyroid. 2007;17(3):271–275.
[267] Zanchetta MB, Giacoia E, Jerkovich F, et al. Asymptomatic elevated parathyroid hormone level due to immunoassay interference. Osteoporos Int. 2021;32(10):2111–2114.
[268] Fiad TM, Duffy J, McKenna TJ. Multiple spuriously abnormal thyroid function indices due to heterophilic antibodies. Clin Endocrinol. 1994;41(3): 391–395.
[269] Lam L, Bagg W, Smith G, et al. Apparent hyperthyroidism caused by biotin-like interference from IgM anti-streptavidin antibodies. Thyroid. 2018;28(8): 1063–1067.
[270] Wouters Y, Oosterbos J, Reynaert N, et al. Alarmed by misleading interference in free T3 and free T4 assays: a new case of anti-streptavidin antibodies. Clin Chem Lab Med. 2020;58(3):e69–e71.
[271] Favresse J, Paridaens H, Pirson N, et al. Massive interference in free T4 and free T3 assays misleading clinical judgment. Clin Chem Lab Med. 2017;55(4): e84–e86.

[272] Sakata S, Matsuda M, Ogawa T, et al. Prevalence of thyroid hormone autoantibodies in healthy subjects. Clin Endocrinol. 1994;41(3):365–370.
[273] Ni J, Long Y, Zhang L, et al. High prevalence of thyroid hormone autoantibody and low rate of thyroid hormone detection interference. J Clin Lab Anal. 2022;36(1):e24124.
[274] Zouwail SA, O'Toole AM, Clark PM, et al. Influence of thyroid hormone autoantibodies on 7 thyroid hormone assays. Clin Chem. 2008;54(5):927–928.
[275] Ylli D, Soldin SJ, Stolze B, et al. Biotin interference in assays for thyroid hormones, thyrotropin and thyroglobulin. Thyroid. 2021;31(8):1160–1170.
[276] Colon PJ, Greene DN. Biotin interference in clinical immunoassays. J Appl Lab Med. 2018;2(6):941–951.
[277] Mzougui S, Favresse J, Soleimani R, et al. Biotin interference: evaluation of a new generation of electrochemiluminescent immunoassays for high-sensitive troponin T and thyroid-stimulating hormone testing. Clin Chem Lab Med. 2020;58(12):2037–2045.
[278] Choi J, Yun SG. Comparison of biotin interference in second- and third-generation Roche free thyroxine immunoassays. Ann Lab Med. 2020;40(3):274–276.
[279] D'Aurizio F, Biasotto A, Cipri C, et al. Thyroid function tests, incongruent internally and with thyroid status, both in a pregnant woman and in her newborn daughter. Eur Thyroid J. 2022;11:e210088.
[280] Ricci V, Esteban MP, Sand G, et al. Interference of anti-streptavidin antibodies: more common than we thought? In relation to six confirmed cases. Clin Biochem. 2021;90:62–65.
[281] Weiss RE, Refetoff S, et al. Thyroid function testing. Chapter 77. In Jameson JL, De Groot JL, de Kretser DM eds. Endocrinology: Adult and pediatric. 7th ed. Philadelphia, PA: Saunders Elsevier; 2016. p. 1444–1492.
[282] García-González E, Aramendía M, Álvarez-Ballano D, et al. Serum sample containing endogenous antibodies interfering with multiple hormone immunoassays. Laboratory strategies to detect interference. Pract Lab Med. 2016;4:1–10.
[283] Mongolu S, Armston AE, Mozley E, et al. Heterophilic antibody interference affecting multiple hormone assays: is it due to rheumatoid factor? Scand J Clin Lab Invest. 2016;76(3):240–242.
[284] Trambas C, Lu Z, Yen T, et al. Depletion of biotin using streptavidin-coated microparticles: a validated solution to the problem of biotin interference in streptavidin-biotin immunoassays. Ann Clin Biochem. 2018;55(2):216–226.
[285] Goettemoeller T, McShane AJ, Rao P. Misleading FT4 and FT3 due to immunoassay interference from autoantibodies. Clin Biochem. 2022;101:16–18.