Contemporary Issues in Thyroid Disease Measurements—Slide Notes

The laboratory support for physicians diagnosing and monitoring thyroid disease necessitates knowledge of both thyroid pathophysiology and the technical limitations of current biochemical tests.

This presentation will address four controversial issues:

- The TSH reference range controversy
- The current imitations of free T4 immunoassay methodology
- Strengths and pitfalls of thyroid autoantibody testing -
- Clinical utility of 2nd generation Thyroglobulin (Tg) measurement

1. The TSH reference range controversy

Slide 3

Currently TSH is the most important thyroid test but there are controversies concerning the setting of the reference range.

Slide 4

In the course of developing hypo- or hyperthyroidism, TSH is the first abnormality to appear as soon as the pituitary registers that free T4 has fallen below its setpoint. The setting of the TSH reference range is thus critical for detecting mild subclinical hypo- or hyperthyroidism.

Slide 5

Over the last four decades, the TSH upper reference limit has contracted nearly 3- fold as a result of technical improvements and more rigorous selection of reference populations. In fact, some guidelines have recommended the adoption of a TSH upper limit of 2.5 to 3.0 mIU/L.

Slide 6

A number of professional organizations have recommended narrowing the TSH upper limit to 2.5 - 3.0 mIU/L - lower than most labs and manufacturers cite. Specifically, AACE has recommended adopting a TSH reference range of 0.3 - 3.0 mIU/L whereas the NACB Guidelines recommend a TSH upper limit of 2.5 mIU/L which is in accord with the new Endocrine Society guidelines for 1st. Trimester pregnancy.

Slide 7

In one study of the reference ranges cited by 30 laboratories using the same method were very different. Three laboratories established their own reference range – an approach that is generally recommended. 30% of laboratories stated they had adopted the manufacturers reference range - but even these ranges differed as if the laboratories couldn't read the kit package insert! The majority of (60%) laboratories claimed they used an "adapted" reference range. These varied and were of unclear origin although closer questioning revealed that these were arbitrarily set. **So even when laboratories use the same TSH method the upper limit on the lab report might vary between 3.8 and 6.0 mIU/L and the lower limit between 0.2 and 0.4 mIU/L!**

Slide 8

When using different manufacturers assays reference ranges are just as variable. Lower limits varied between 0.27 and 0.49 and upper limits between 3.1 and 5.5 mIU/L. When the same specimen was measured in all the assays the TSH reported varied between 1.07 and 0.62 mIU/L - differences that were not related to differences in the reference range.

Slide 9

Why these differences in TSH reference limits?

Guidelines recommend that TSH reference limits be calculated from cohorts of subjects judged to be euthyroid.

However these ranges are influenced by a myriad of factors such as the population demographics – ethnicity and age (both of which influences the prevalence of thyroid autoimmunity) and iodine intake.

Slide 10

The NHANES III survey found that African Americans have lower TSH values than Mexican Americans or Caucasians. In fact, the upper TSH limit for African Americans was 3.6 compared to 4.1 and 4.2 for the other groups.

When this same database was analyzed relative to age a shift in TSH towards higher values for the older age groups was seen. For example, the TSH upper limit for 20-29 year olds was 3.5 mIU/L compared with 7.5 mIU/L for over 80 year olds.

Slide 11

TSH reference limits are also influenced by the rigor used to exclude subjects with risk factors for thyroid dysfunction – family history, thyroid autoimmunity as evident from the presence of thyroid antibodies or a hypoechogenic ultrasound.

Slide 12

When the NHANES III database analyzed the relationship between TSH and the prevalence of thyroid antibodies the lowest antibody prevalence was found to be in the range of 0.1 to 2.0 mIU/L for women and 0.1 to 1.5 mIU/L for men. Most antibodies were associated with TSH levels above these cutoffs. However, even when TSH is unequivocally elevated (>20 mIU/L) **not all individuals were antibody-positive**. So the inclusion of individuals with mild thyroid dysfunction but without detectable antibodies will tend to skew the TSH upper limit.

Slide 13 &14

One thing rarely considered is that assay specificity differs. Different assays detect different circulating TSH isoforms, some of which may not be biologically inactive. An extreme example is the paradoxically normal TSH levels typical of central hypothyroidism. This arises because current assays measure the biologically inert TSH isoforms that are secreted in cases of pituitary failure.

Slide 15

In the Kratzsch study (Clin Chem 51:1480, 2005) the same cohort of specimens from rigorously selected healthy blood donors gave a TSH upper limit of 3.8 when using the Elecsys assay but an upper limit of only 2.9 mIU/L when the Bayer assay was used to measure TSH in the same sera.

Slide 16

Comparing TSH reference ranges across different studies you see a remarkable consistency in the TSH lower limit (0.3 - 0.4 mIU/L) but a TSH upper limit that is clearly influenced by geography - through iodine deficiency, ethnicity -through differences in the prevalence of autoimmune thyroid disease and TSH assays - that detect different TSH isoforms. So the TSH upper limit is a moving target and has been reported to lie anywhere between 4.2 and 2.4 mIU/L depending on the combination of these factors.

Slide 17

Graphically, the lowest TSH upper limit reported is a 2.1 mIu/L from the Ship German study, fairly close to the 2.5 recommended by the Endocrine Society for 1st. Trim pregnancy. AACE guidelines and Endocrine Society Guidelines for 2nd & 3rd trimester pregnancy recommend an upper limit of 3.0 mIU/L. The NHANES SURVEY reported a TSH upper limit of 3.6 mIU/L for 20-29 year olds and the African American population but higher values around 4.0 mIU/L for Mexican Americans, 50-59 year olds and Caucasions. The highest TSH upper limit is reported as 7.5 mIU/L for over 80 year olds.

Slide 18 & 19

The other important issue that should impact the use of the TSH reference range is the fact that thyroid tests, whether considering TSH, total or free thyroid hormones or thyroglobulin (Tg) have narrow within-person variability relative to the between-person variability - That is they have a low indexes of individuality (IoI). This means that even values within the reference range can be abnormal for an individual patient because a considerable degree of abnormality has to develop in the individual before a thyroid test result moves outside its reference range.

Slide 20

In a study in which TSH was measured every month for a year in euthyroid control subjects it is clear that TSH in individuals varies over a narrow range (1.0 mIU/L) -much narrower than the span of the typoical population reference range. So as an individual begins to develop thyroid failure TSH would rise to a value that was clearly abnormal for that individual but was still within the population reference range.

The current controversy concerns:

- Would there be any consequences for that patient if left untreated?
- Should that patient be classified as subclinically hypothyroid if the TSH is still within the population reference range?
- At what TSH level should replacement L-T4 be considered?

Slide 21

There is growing awareness that clinical stratification is more important than the accuracy of the setting of the TSH upper reference limit. The comprehensive review by Drs Biondi and Cooper (Endoc Rev 29:76, 2008) rationalizes that when TSH appears elevated (but is still below 10 mIU/L) patient-specific factors related to cardiovascular risk such as symptoms, positive TPOAb, pregnancy or infertility should guide the need for L-T4 replacement on a case-by case basis.

Slide 22
Current guidelines recommend that TSH ranges be patient-specific:
• Laboratory reference ranges are variable yet an individual's TSH only varies +/- 0.5 mIU/L over time.
• The TSH upper limit for pregnancy and preconception is similar to that for L-T4 replacement for hypothyroidism.
• Because TSH is a trophic factor for thyroid tissue TSH targets for thyroid cancers are patient-specific and related to risk factors for persistent/recurrent disease.

2. Limitations of free T4 (FT4) immunoassay methodology

Slide 24
Most (> 99%) thyroid hormone in blood is bound to plasma proteins, mainly TBG (and in the case of T4 also albumin, transthyretin and some other proteins). Changes in binding proteins can influence TT4 independent of the minute free T4 (FT4) fraction. However it is only FT4 that is able to enter peripheral tissues and be converted to the more metabolically active thyroid hormone T3 that exerts biologic activity.

Slide 25
When physicians check off Free T4 on the test requisition that are ordering a FT4 estimate test because direct FT4 measurements (those that physically separate free from bound hormone) are typically only available in reference labs. FT4 estimates are either based on a FT4 index approach (a correction of TT4 for the TBG concentration) or a FT4 immunoassay. FT4 immunoassays are designed to measure FT4 without separating free from bound hormone are based on either a one-step, two-step or labeled antibody approach.

Slide 26
Studies have shown that current FT4 immunoassays are binding protein dependent.
In the study by Fritz et al (Clin Chem 53:911, 2007) when protein-bound TT4 was held constant and FT4 measured by direct equilibrium dialysis was increased across 4 logs of concentration, FT4 immunoassay measurements did not and did not detect the rise in FT4 measured by equilibrium dialysis. Conversely, when FT4 was kept constant and serum proteins and were varied to change TT4, again FT4 immunoassays did not reflect FT4 measured directly by mirrored the increasing TT4 concentration.

Slide 27
Current FT4 immunoassay methodology has been specifically engineered to overcome TBG abnormalities that are commonly encountered. However, as a result these tests include proprietary blockers and binders that make them subject to albumin and sensitivity. This albumin sensitivity decreases the diagnostic accuracy of these tests used to assess thyroid status during pregnancy, nonthyroidal illness and congenital albumin abnormalities (Familial Dysalbuminemic Hyperthyroxinemia, FDH syndromes).

Slide 28
Low albumin is common in conditions of nonthyroidal illness (NTI) in which there is a well-described spectrum of changes in thyroid tests as severity increases and during recovery. As the

severity of illness increases, discordance develops between FT4 measured by direct equilibrium dialysis that remains relatively normal (as appears appropriate for the euthyroid status of such patients) versus TT4 and FT4 immunoassay measurements that become progressively lower.

Slide 29

The FT4 immunoassay values in hospitalized patients with NTI are likely spuriously low because of the low albumin state of such patients. In vitro studies have shown that adding albumin to low albumin NTI sera can normalize the apparent FT4, whereas adding albumin had no effect on FT4 measured by equilibrium dialysis.

Slide 30

Pregnancy is a common condition in which there are significant binding protein abnormalities that impact the reliability of FT4 immunoassay measurements:

• First the estrogen-mediated rise in TBG that is responsible for the 50% increase in TT4.

• hCG shares some structural homology with TSH and the hCG peak of the 1st trimester directly stimulates the thyroid to produce a transient rise in FT4 which can be detected when measured directly by equilibrium dialysis/ tandem mass spectrometry and a reciprocal fall in TSH.

• FT4 immunossays fail to show the early hCG-mediated rise in FT4 and progressively fall as gestation progresses.

• the declining FT4 immunoassay values likely reflect the progressive decline in albumin (~15-20% by the 3rd trimester).

Slide 31

The study of Sapin et al (Clin Lab 50:581, 2004) showed that when using different FT4 immunoassays the percentage of patients with FT4 below the non-pregnant reference range by the 3rd trimester ranged from 62% 14%. This was clearly a method-dependent decline likely reflecting the albumin-dependancy of the method and not physiology.

Slide 32

When trimester related changes are expressed relative to a % of the non-pregnant mean, FT4I corrects the high TBG effect on TT4. Both FT4I and FT4 measured directly by equilibrium dialysis and tandem mass spectrometry show the hCG-mediated 1st trimester rise, but usually remain within the non-pregnant reference range throughout gestation. In contrast, FT4 immunoassays often fail to show the first trimester peak and progressively decline to a method-related degree throughout gestation, so that by the 3rd trimester a significant proportion of the pregnant patients have FT4 values reported below the non-pregnant range. This is likely to reflect a low albumin artifact.

Slide 33

The new Endocrine Society guidelines for pregnancy stress the importance of keeping TSH below 2.5 mIU/L especially in the 1st trimester, but state that when FT4 measurements are needed such during anti-thyroid drug treatment for hyperthyroidism, the FT4 target should be the upper third of the non-pregnant reference range for that assay.

3. • Thyroid autoantibody testing - strengths and pitfalls

a) TSH Receptor antibodies (TRAb):
b) TPOAb, Thyroid Peroxidase antibodies
c) Thyroglobulin antibodies (TgAb)

Slide 35
Most manufacturers standardize their methods against WHO reference preparations and yet the cut-offs for positivity are highly variable which is why the values of different manufacturers tests cannot be compared. This is more of an issue with TgAb tests because serial TgAb is used as a tumor marker for DTC.

a) TSH Receptor antibodies (TRAb):

Slide 36
There are two classes of TSH receptor antibodies:
- stimulating antibodies that cause Graves' hyperthyroidism and
- blocking antibodies that block the action of TSH are rare but can cause hypothyroidism

Slide 37
Measurement of these antibodies is rarely needed because the diagnosis of Graves' hyperthyroidism in most cases is evident from the clinical presentation, especially the characteristic eye signs of diplopia, eye discomfort and protrusion.

Slide 38
TRAb measurements are primarily used to assess risk of neonatal hyper- or hypothyroidism in the 3rd trimester when mothers have a history of Graves' disease. Maternal and fetal thyroid function are controlled separately. However, the there is transplacental passage of thyroxine, anti-thyroid drugs and antibodies including TRAb. When high levels of TRAb cross the placenta they have the potential to stimulate or block fetal thyroid function, depending on whether the antibody stimulates or blocks the action of TSH.

Slide 39
TRAb tests are not useful for predicting Graves' disease remission after medical treatment or predicting severity and outcome of Graves' opthalmopathy. TRAb testing can be useful for assessing neonatal risks when pregnant mothers have active Graves' or have had prior surgical or radioiodine treatment for Graves' hyperthyroidism, for determining the etiology of hyperthyroidism and for diagnosing euthyroid Graves' opthalmopathy.

b) **Thyroid Peroxidase antibodies, TPOAb**

Slide 41
TPOAb is the most sensitive marker for thyroid autoimmunity as evident from the NHANES III survey that found that TPOAb is strongly associated with overt and subclinical hypothyroidism. 14.7% of the US population had either TPOAb or TgAb detected, the dominant antibody being TPOAb. It was significant that the odds ratios for overt and subclinical hypothyroidism were strongly associated with the presence of TPOAb, either associated with TgAb or alone, but not with TgAb alone.

Slide 42
The strong association between TPOAb and thyroid deficiency is not surprising because TPOAb appears to mediate antibody-dependent and complement-dependent cytotoxicities. Studies have suggested that TPO antibodies are cytotoxic to thyroid cells.

Slide 43
However, TPO antibodies are not always present in all affected individuals developing autoimmune thyroid dysfunction. In the early phases of developing thyroid failure a hypoechoic ultrasound pattern indicative of lymphocytic infiltration may be the only indicator of the condition apart from a slightly elevated TSH.

Slide 44
TPOAb has a number of clinical uses. TPOAb is primarily used to diagnose or as a risk factor for autoimmune thyroid disease. It is risk factor for a the development of thyroid dysfunction especially following the initiation of interferon-alpha, interleukin-2, lithium or amiodarone therapy and in Down's syndrome patients In the reproductive age women the presence of TPOAb is a risk factor for, miscarriage, IVF failure and post-partum thyroiditis.

c) Thyroglobulin antibodies (TgAb)

Slide 45
TgAb measurements are secondary tests for diagnosing autoimmune thyroid disease but the primarily test used to assess risk for TgAb interference with Tg measurements made for patients with differentiated thyroid cancer (DTC). Unfortunately the prevalence of TgAb in DTC is twice that of the general population.

Slide 46
TgAb interference with serum Tg measurements is a serious problem. Most laboratories used Immunometric Assay (IMA) methods that are prone to TgAb and heterophilic antibody (HAMA) interferences whereas RIA methodology is resistant to such interferences. Clinical laboratories favor IMA methods over the older RIA methods because IMAs are faster and can be automated.

Slide 47
When TgAb is present Tg circulates free or bound to the endogenous antibody. It appears that RIA methods are able to quantify both free and antibody-bound Tg, whereas IMA measurements are often only able to measure free Tg. This is why falsely low or undetectable Tg IMA measurements are frequently reported when TgAb is present.

Slide 48
All Tg IMAs are subject to report falsely low Tg measurements when TgAb is present. This was evident from measurements made in TgAb-positive euthyroid control subjects with intact thyroid glands. Each of four different IMA methods reported some paradoxically undetectable or low values for a significant number of these subjects, all of whom should have had some Tg detected. In contrast, Tg was appropriately within the reference range established for TgAb-negative controls when RIA methodology was used.

Slide 49
Reporting falsely low Tg values for DTC patients is clinically unacceptable because studies show that patients with persistent TgAb have a higher stage of disease and more risk of lymph node metastatases. However, when a laboratory does not have access to a Tg RIA method, changing TgAb concentrations can be used as a surrogate tumor marker.
For example when TgAb declined more than 50% 6-12 months after surgery no patients showed evidence of persistent disease as compared with 37% in whom TgAb increased.

Slide 50
Unfortunately current TgAb methods do not always detect interfering TgAb. In one study (Spencer et al JCEM 90:5566, 2005) TgAb was detected by one or more methods in 42 specimens. The degree of positivity varied from 8% (one method +) to 100% (all methods positive). Only 3/42 specimens were reported as positive by all methods. Some methods did not detect TgAb in some specimens and the absolute numbers reported by each method are different.
TgAb concentrations cannot be used as a surrogate tumor marker unless the measurements are made by the same method.

4. Clinical utility of 2nd generation Thyroglobulin (Tg) measurement

Slide 52
Three pathophysiological factors influence the serum Tg concentration – mass of thyroid tissue, any thyroid injury and the degree of TSH receptor stimulation. In addition it is important to consider the Tg assay characteristics especially functional sensitivity.

Slide 53
As sensitivity becomes increasingly recognized as critically impacting Tg assay quality, it is important to avoid the problems experiences with TSH sensitivity improvements in the 1990s where marketing pressures resulted in a plethora of meaningless descriptive terms like “ultrasensitive” and “supersensitive” that did not help laboratories distinguish between the sensitivity of different manufacturers products.

Slide 54
There is a detailed protocol for determining Tg assay functional sensitivity (FS).
Specifically, FS is the Tg value that can be measured with 20% between-run CV (*using 1:1 CRM-457 standardization). The protocol contains some important provisos:*
• Precision should be determined in human serum pools containing no evidence of Tg antibodies (TgAb).
• Precision should be determined across a time-span that relates to the use of the measurement in clinical practice (6 - 12 months)
• This time-span should involve the use of more than 2 reagent lots and 2 instrument calibrations, because one major problem that compromises precision is the consistency of the lots produced by the manufacturer.

Slide 55

A generational nomenclature can be adopted to describe Tg assay functional sensitivity, analogous to TSH. Most current assays are only first generation, with FS between 0.5 and 1.0 ng/mL. Some current assays are 2nd generation with 10-fold better FS (0.05-0.1 ug/L). In the future we will likely have 3rd. generation assays with functional sensitivity-between 0.05 and 0.005 ug/L.

Slide 56

Whether Tg is detectable really depends on the sensitivity of the assay used. Current Tg assays can have 10 fold differences in FS ranging from 2.0 to 0.1 ng/mL. Assays that can only measure down to 1.0 ng/mL are barely able to detect Tg in all normal euthyroid subjects with intact thyroid glands and have inadequate sensitivity for detecting tumor in thyroidectomized patients!

Slide 57

Tg assay sensitivity has practical and economic consequences. It has become common practice to use recombinant human TSH (rhTSH) stimulation to overcome the insensitivity of the 1st generation assays. Patient with an undetectable Tg (<1.0 ng/mL) sometimes have a positive rhTSH stimulated Tg response above 2.0 ng/mL and are at higher risk of having disease. When a 2nd generation assay is used such patients will often have a detectable basal Tg in the 0.1 – 1.0 ng/mL range. Basal Tg is typically ~8 fold lower than rhTSH stimulated Tg.

Slide 58

Currently, rhTSH stimulation is used to compensate for Tg assay insensitivity. When using a 1st generation assay three rhTSH responses are seen:

• Patients with a detectable basal Tg typically display a 8-fold increase in Tg above basal in response to rTSH and disease is detected in about half. In this group, rhTSH stimulation adds no additional value.

• 80% of patients with Tg below 1ng/ml display no rhTSH stimulation have a low risk of having disease.

• 20% of patients with basal Tg below 1.0 ng/mL have rhTSH above 2.0 ng/mL and ~6% of these patients have disease.

Slide 59

Using a more sensitive 2nd generation assay will detect responses that were undetectable using a 1st generation assay. The approximate 8-fold relationship between basal and rhTSH-stimulated Tg is retained even when basal Tg is below 1.0 ng/mL.

Slide 60

Basal Tg correlates with rhTSH-stimulated Tg, even when Tg levels are below 1.0 ng/mL. In one study (Spencer et al Nat Clin Pract Endocr Metab 4:223, 2008) only 1/373 patients with a basal Tg below 0.1 ng/mL had a positive rhTSH stimulated Tg response above 2 ng/mL and only 4 patients had a rhTSH-stimulated Tg value between 1 and 2 ng/mL.

Slide 61

Tg assay insensitivity has practical and economic consequences. When basal Tg is measured below 0.1 ng/mL by a 2nd generation assay it is most unlikely that you will find a rhTSH stimulated Tg above 2 ng/mL. There is growing consensus that rhTSH is unnecessary when a 2nd

generation assay is used. Thus a switch to more a sensitive 2nd generation Tg assay maximizes the clinical utility of serial Tg testing and is likely to make rhTSH stimulated Tg testing obsolete, analogous to the decline in TRH stimulation testing once 3rd. generation TSH assays became available.

To conclude

Slide 62

1. **TSH reference range controversy**

The TSH upper reference limit is assay dependent and less important than evaluating TSH status relative to the patient's risk factors for CVD.

2. ***Limitations of free T4 immunoassay methodology***

FT4 immunoassays binding protein (albumin) dependent.
Total T4, free T4 indexes (FT4I) and FT4 reference methods (ED+TMS) are preferred when evaluating FT4 status in pregnancy and NTI.

Slide 63

3. Thyroid autoantibody testing - strengths and pitfalls

Cannot compare absolute values reported by different methods
TRAb: to assess Graves' mothers for neonatal risks
TPOAb: primary marker for thyroid autoimmunity
TgAb: surrogate DTC tumor marker & risk for Tg interference

4. 2nd generation Thyroglobulin (Tg) measurement

• Tg assay variability precludes switching methods
• 2nd generation assays (functional sensitivity ≤ 0.1 ng/mL)
obviate the need for expensive rhTSH stimulated Tg