Age- and Race-Based Serum Thyrotropin Reference Limits

Martin I. Surks and Laura Boucai

Division of Endocrinology, Department of Medicine (M.I.S., L.B.) and Pathology (M.I.S.), Montefiore Medical Center and the Albert Einstein College of Medicine, Bronx, New York 10467

Context: TSH reference limits, particularly the upper limit, are controversial. The traditional and prevailing method for setting limits uses TSH distribution of thyroid disease-free individuals. The curve is not Gaussian, but skewed to higher concentrations, even after log-transformation; values in the skewed area are assumed to reflect mild hypothyroidism. The underlying assumption for this traditional approach, which has not previously been tested, is that the limits derived from this curve are applicable to all people. However, recent studies suggest that distinct subpopulations have unique TSH distribution and reference limits that are significantly different from limits established by the traditional approach.

Evidence Acquisition: A search was focused on articles that provide the basis for current recommendations for setting TSH reference limits as well as articles that suggest that the traditional method does not reflect accurately the TSH distribution and reference limits of distinct subpopulations within the United States.

Evidence Synthesis: TSH distribution and reference limits shift to higher concentrations with age, even up to centenarians, and are unique for different racial/ethnic groups, being at higher concentrations in Caucasians than either Blacks or Hispanics originating from Puerto Rico or the Dominican Republic. The distribution curve derived by the traditional approach represents a composite of curves from specific subpopulations that do not provide appropriate reference limits for those unique groups.

Conclusions: Age- and race-specific TSH distribution and reference limits, possibly influenced by genetic factors, should be employed to provide clinicians accurate limits for specific populations and guidance for further evaluation of thyroid dysfunction. (J Clin Endocrinol Metab 95: 496–502, 2010)

Serum TSH is the most sensitive index of thyroid function in the absence of hypothalamic-pituitary disease. During the last 5 yr, several events have focused attention on the determination of TSH reference limits, particularly the upper limit. First, a detailed analysis of TSH, T4 and anti-thyroid antibodies in a representative sample of the U.S. population [National Health and Nutrition Examination Survey III (NHANES III)] (1) showed a dramatic and progressive increase in prevalence of TSH above 4.5 mIU/liter with age, to up to 15% of individuals more than 70 yr old without thyroid disease or risk factors for thyroid disease. Because about 35 million Americans are currently older than 70 yr of age, 5.25 million would be predicted to have raised TSH and therefore be designated hypothyroid by that analysis; 85% of those with TSH above 4.5 mIU/liter had normal serum T4, and would therefore be diagnosed with subclinical hypothyroidism. Second, in 2003, the American Association of Clinical Endocrinologists abruptly recommended a decrease in the upper limit of TSH to 3.0 mIU/liter (2); other experts recommended a further decrease to 2.5 mIU/liter (3, 4), possibly based on a similar suggestion in a monograph on

Abbreviations: AB, Antithyroid antibodies; SNP, single nucleotide polymorphism.
laboratory medicine practice guidelines by the National Academy of Clinical Biochemistry (NACB) (5). However, these recommendations were controversial (6, 7) and still remain unsupported by studies of TSH distribution in thyroid disease-free populations. Moreover, if those lower upper limits had been widely accepted, the prevalence of raised TSH would have increased nearly 4-fold, designating about 20 million Americans with raised TSH and possible hypothyroidism (6, 7). Lastly, the ongoing debate over subclinical hypothyroidism, particularly its definition, association with adverse health outcomes, recommendations for screening, and use of levothyroxine for treatment (8–10) has maintained focus on defining the TSH upper limit.

Genetic Influence on Thyroid Hormone and TSH Concentrations

In recent years, a number of reports have indicated hereditary and genetic influences on concentrations of free T$_4$ and TSH. In an analysis of 12 monthly measurements of free T$_4$ and TSH in 16 healthy men, Andersen et al. (11) showed that the intraindividual variation in these hormone concentrations was far smaller than the interindividual variation, demonstrating a high degree of reproducibility in each individual’s measurements over time. These findings suggested possible hereditary influences on the concentrations of free T$_4$ and TSH and the negative feedback set point. This thesis is supported by studies in identical and fraternal twins that show a strong effect of genetic factors and heredity on thyroid hormone and TSH concentrations (12, 13) as well as the negative feedback set point. Furthermore, a large study of Mexican-American families, the San Antonio Family Heart Study (14), suggested that genes controlled a substantial portion of the variation in thyroid hormones. Lastly, recent studies (15–19) showed that specific polymorphisms in genes that code for the thyroid hormone receptors, the iodothyronine deiodinases, and the TSH receptor are associated with unique changes in serum thyroid hormone and TSH concentrations.

Even if each individual has a unique genetically determined set point of serum thyroid hormones and TSH, those previous concentrations are rarely available to clinicians for comparison to current hormone measurements. Thus, genetically determined hormone concentrations are of limited practical help for diagnosis of thyroid dysfunction. Practitioners must therefore depend on thyroid hormone and TSH distributions and reference limits derived from population studies to determine whether their patients fall within those limits or not. Such limits may include TSH concentrations from a small number of people with thyroid disease whose genetically determined serum TSH has changed, but not sufficiently to exceed population-based reference limits. In fact, based on NHANES III data (20), approximately 13% of the U.S. population without known thyroid disease has antithyroid antibodies, and the majority of these individuals (nearly 60%) have serum TSH between 0.4 and 2.5 mIU/liter.

At present, there are no data to support defining thyroid dysfunction and providing treatment to patients whose serum TSH concentration either minimally exceeds population-based upper limits, the 97.5 centile, or is minimally below the lower limit, 2.5 centile. Treatment is generally recommended if serum TSH consistently exceeds 10 mIU/liter or remains less than 0.1 mIU/liter (8, 10). When the serum TSH is at or minimally above or below population-based limits, usually designated subclinical thyroid dysfunction, the clinician should be alerted to assess the patient carefully for thyroid disease. Treatment decisions for individual patients should be based on the concentration of serum TSH, thorough clinical evaluation, and within the framework of studies concerning adverse health outcomes that are summarized in recent guidelines and reviews (8, 10).

In this article, we examine the basic precepts and assumptions underlying the prevailing interpretation of the TSH distribution curve and methods for setting reference limits for populations. Secondly, we review new findings, which suggest that the prevailing interpretation does not accurately reflect TSH distribution in specific age, race, and ethnic groups, and finally, we evaluate the clinical implications of using one distribution curve for all populations. We do not discuss the particular case of pregnant women because this subject has recently been reviewed (21).

Prevailing Method for Determination of TSH Reference Limits

Development of the RIA for TSH by Utiger (22) was a landmark in endocrinology, subsequently enhanced by employing monoclonal antibodies and immunometric analysis, which resulted in a dramatic increase in assay sensitivity. Clinicians can now readily distinguish undetectable TSH in patients with hyperthyroidism from values of normal individuals. However, differences in interlaboratory and intermethod precision, functional sensitivity, particularly for decreased or suppressed TSH, as well as assay specificity became apparent (23–25). Because many institutions and independent laboratories perform measurements of TSH using reagents from a number of sources, concerns about assay reliability persist.
The recognition of these problematic issues, among others, led to the development of rigorous standards for TSH determination. In 2002, the NACB published consensus guidelines that provided methodology and criteria for optimal TSH determination, including quantitation of assay interference, functional sensitivity, between-run precision, and determination of reference limits (5). They suggested that reference limits be established from TSH concentrations of at least 120 individuals who do not have thyroid disease. Such individuals should not have visible goiter, family or personal history of thyroid disease, take medications that can affect thyroid function, or have antithyroid antibodies (AB). The TSH distribution curve of such individuals is not Gaussian but is right-skewed toward higher TSH concentrations, even after log-transformation (5). Recommended analyses employ nonparametric statistics and define the reference range using the median TSH and 95% confidence limits. Despite small but statistically significant differences in median TSH and reference limits in different genders, races, and ages reported in the NHANES III analysis, the authors suggested that TSH reference intervals need not be individually adjusted for these populations (1).

Does One TSH Distribution Curve Fit All?

The prevailing basis for interpreting TSH reference limits for populations is that one distribution curve, developed as recommended in the NACB guidelines, can be used for all people in the U.S. population; that is “one size fits all.” Two principal matters of debate are the inclusion of patients with thyroid ultrasound abnormalities and the implications of the right skew or “tail” in the TSH distribution curve.

A minority of patients with undetected autoimmune thyroid disease may have specific thyroid ultrasound abnormalities, including hypochoegenicity and heterogeneity of the echo signal, even in the absence of AB (26). However, in one analysis, more than 70% of AB− subjects who had both hypochoegenicity and “irregular echo” had TSH less than 3.6 mIU/liter (27), suggesting that the majority of individuals with these ultrasound abnormalities do not have hypothyroidism.

The right skew of the TSH distribution curve has received much attention and has been assumed to reflect inclusion of patients with thyroid disease and early thyroid failure (5). If that were true, the descending limb of the distribution curve could be extrapolated to the x-axis, with an intercept of about 2.5 to 3.0 mIU/liter, as suggested (4, 5), and those individuals with TSH above that limit could be considered hypothyroid. The evidence supporting this assumption is a single report that showed a significant increase in rate of progression to overt hypothyroidism when TSH was above 2.0 mIU/liter (28). For all patients with TSH above 6.0 mIU/liter initially, the 20-yr incidence of overt hypothyroidism was 33% in the absence of antithyroid antibodies (AB−) and 55% in the presence of antithyroid antibodies (AB+). Extrapolation of a curve relating the probability of developing hypothyroidism to the initial serum TSH was significant statistically even when the initial TSH was in the 2.0 to 5.0 mIU/liter range (28).

However, a recent reanalysis of these data showed that progression to overt hypothyroidism was much slower when the initial TSH was 2.0 to 5.0 mIU/liter (7); 10–15% per 20 yr in AB− and 20–40% per 20 yr in AB+ patients. An important context for these findings is that even when initial serum TSH was 1–2 mIU/liter, overt hypothyroidism developed in 4% per 20 yr in AB− individuals and 20% per 20 yr in AB+ individuals (7). Thus, although significant statistically, it is likely that only a minority of people with TSH between 2 and 5 mIU/liter have thyroid disease. This conclusion is generally supported by the relatively low prevalence of AB (23.9%) when TSH was 2.5–4.5 mIU/liter in clinically thyroid disease-free subjects reported in NHANES III (1, 20).

The assumption that patients with autoimmune thyroid disease with mild hypothyroidism account for the skew in TSH distribution curves has now been tested. Three reports demonstrate that median TSH and reference limits were not significantly influenced by exclusion of patients with AB (29) or patients with both AB and ultrasound abnormalities (30, 31). These findings support the conclusion that only a minority of people whose TSH falls in the skew of the TSH distribution curve have thyroid disease and that most do not. A recent review (7) of studies of possible reasons that patients’ TSH might fall in the skew postulated that the majority of individuals, who do not have thyroid disease, may have an exaggerated diurnal variation or a shift in circadian TSH secretion caused by working either the night shift or swing shift, or going to sleep very late, engaging in vigorous exercise, having rare congenital mutations in the TSH receptor, having TSH with decreased biological activity, or being obese. Other studies also show minimal increase in TSH in patients who suffer from depression (32). None of these putative explanations for the skew has been examined in population studies.

One TSH Distribution Curve Does Not Fit All

An alternative explanation for the skewed TSH distribution curve is that the currently employed curve is in fact a
A composite of unique curves for specific subpopulations, such as people of different ages and ethnicities which, when blended into a single distribution curve, creates the appearance of a skew. If true, TSH reference limits derived from the currently employed composite standard curve would not appropriately apply to these subpopulations and could lead to incorrectly designating patients either within or outside of their respective population-specific reference limits. For example, the NHANES III report described a progressive increase in median TSH and reference limits with age for the thyroid disease-free population, even after excluding those with AB (1). It also showed that median TSH and reference limits were lower in Black, non-Hispanics compared with White, non-Hispanic individuals. The possibility that these changes in median TSH and reference limits reflect TSH distributions unique for these populations has recently been studied (20, 33, 34).

Theoretically, if a shift in median TSH derived from a composite distribution curve reflects inclusion of patients with thyroid dysfunction, the distribution curve would have lower amplitude at peak frequency TSH concentration, which remains unchanged, and a significant skew would be apparent (Fig. 1A). In contrast, when a shift in median TSH represents the normal TSH distribution of a subpopulation, the entire distribution curve would be displaced in the direction of the changed median, including the TSH concentration at peak frequency (Fig. 1B).

**Age-Specific TSH Distribution and Reference Limits**

Using the NHANES III and NHANES 1999–2002 databases, TSH distribution curves for thyroid disease-free subjects appeared unique for different age groups and shifted progressively to higher TSH concentrations with age (Fig. 2) (20). Similar age-dependent shifts in TSH distribution were reported in an urban outpatient practice of medicine, for well-defined Black, Caucasian, and Hispanic subgroups (33). The shift in TSH distribution to higher concentrations with age has now been extended to individuals who have achieved extreme longevity (34). TSH distribution of a homogeneous cohort of Ashkenazi Jewish centenarians (median age, 98 yr) was compared with unrelated Ashkenazi controls, (median age, 69 yr) and to age-matched controls from NHANES III (median age, 70 yr). TSH distributed at higher concentrations in the centenarians than in either control group (34). In another study (19), TSH distribution of centenarians was also shifted to significantly higher concentrations than their offspring, and that of their offspring was at higher TSH concentrations than their spouses, who served as age-matched controls. These four reports suggest that the increase in median TSH with age reflects mainly population shifts in TSH distribution and support the use of age-specific reference limits in clinical practice.

Additionally, significant heritability of TSH between the Ashkenazi Jewish centenarians and their offspring was observed and was associated with two single nucleotide polymorphisms (SNPs) in the TSH receptor gene (19). The prevalence of the two SNPs was higher in the centenarians and offspring than in the offspring’s spouses, and, within groups, those who had these SNPs had higher TSH concentrations than those who did not. If these findings can be extended to the other populations that demonstrate the age shift in TSH distribution, they would suggest a genetic origin for this change that could provide a basis for the increased prevalence of raised TSH with aging. Moreover, the increased prevalence of these SNPs in centenarians raises the interesting hypothesis that the presence of the two TSH receptor SNPs, associated with increased TSH, may contribute to healthy aging (19).
Race-Specific TSH Distribution and Reference Limits

Several studies suggest that serum TSH in Blacks is lower than in Caucasians (35–37), and the NHANES III report (1) showed a significant decrease in median TSH in Blacks compared with Caucasians: 1.14 mIU/liter compared with 1.43 mIU/liter, respectively. Because TSH distribution curves for the groups were not reported, it was uncertain whether the shift in median TSH resulted from skewed curves or from population shifts in TSH distribution. When TSH distribution curves for Whites, Blacks, and Hispanics in a large urban outpatient medicine practice were analyzed, the curves for Blacks and Hispanics were superimposable. The curves for both of these populations were significantly shifted to lower TSH concentrations than the curve for Caucasians, including TSH concentration at peak frequency (33) (Fig. 3). The ascending and descending limbs of both curves appeared parallel, without significant skew. The median and 97.5th centile for TSH in Whites were significantly higher than in Blacks. Moreover, population shifts in TSH distribution to higher values with aging occurred within each racial group. Median TSH and 97.5th centile were significantly higher in older (>80 yr old) compared with younger (20–29 yr old) Whites, and in older compared with younger Blacks.

A recent analysis (unpublished observations from the NHANES III reference population) showed that TSH distribution and reference limits of Whites and Mexican-Americans were superimposable and that both were significantly shifted to higher TSH concentrations compared with Blacks. The differences between Blacks and Caucasians in NHANES III were almost identical to the comparable populations from an urban outpatient medical practice (33) (Fig. 3). However, TSH distribution of Mexican-Americans in NHANES III, similar to Caucasians, was quite different from that of urban Hispanics in Bronx, New York (33). Because the Hispanic population studied in the Bronx is derived mainly from Puerto Rico and the Dominican Republic, genetic changes between these groups and Mexican-Americans, both of Hispanic origins, may be responsible for the different TSH distribution.

Potential Reclassification of Patients with Thyroid Dysfunction

Given the evidence described above that the population TSH distribution progressively increases with age and shifts with race, significant misclassification of patients with abnormal TSH and possible thyroid dysfunction can occur unless age- and race-specific limits are employed. If the traditional single composite curve from the reference population of NHANES III is used, up to 15% of people more than 70 yr of age exceed the upper reference limit of 4.12 mIU/liter (33). However, a published analysis shows that up to 70% of those considered to have raised TSH actually have values within their age-specific 97.5th centile (20). Using the 2.5th percentile from the NHANES III reference population, i.e. 0.45 mIU/liter, 8% of Blacks and 3.7% of Whites would be misclassified (based on race-specific reference limits) as having decreased TSH. Misclassification may occur even when limits for subpopulations are employed, depending on the thyroid disease-free
volunteers who provide serum for TSH measurement. For example, using the 97.5th centile from curves developed with serum samples from young Blacks, 8% of old Blacks and 10.7% of old Whites would be misclassified with raised TSH and could be considered hypothyroid (33).

**Conclusion, Recommendations, and Challenges**

First, the traditional “one size fits all” method for determining TSH reference limits does not reflect TSH distribution and reference limits of specific ethnic groups and people of different ages. The traditional curve is an integration of unique curves for each of these subpopulations. Some of those curves, especially those of people older than 70 yr of age whose TSH distribution is shifted to higher concentrations, likely account for much of the skew observed in the traditional composite TSH distribution curve. Based on the prevalence of AB, a minority of people with TSH between 2.5 and 4.5 mIU/liter appear to have thyroid disease.

Second, the shift to higher TSH concentration and reference limits with age has a genetic basis in an Ashkenazi Jewish population, associated with the presence of two SNPs in the regulatory/enhancer region of the TSH receptor gene. The prevalence of these SNPs and increased TSH is highest in centenarians and their offspring, compared with controls that are age-matched to the offspring. All populations studied to date exhibit the age-related shift in TSH distribution, including other Caucasians, Blacks, and Hispanics from Puerto Rico and the Dominican Republic, and Mexican-Americans. If these or similar SNPs are related to increased TSH in these elderly populations, genetic changes may be considered beneficial to healthy aging, and even to achieving extreme longevity.

And third, these new findings strongly argue for the use of age- and ethnic/race-specific reference limits for TSH to avoid significant misclassification of patients with abnormal TSH who may or may not have thyroid dysfunction, which occurs when the traditional composite curve and derived limits are employed. This recommendation provides a challenge to clinical pathologists to devise new paradigms for development of age- and ethnic/race-specific TSH reference limits for use in clinical practice. When a patient has serum TSH that is consistently outside of their population-specific reference limits, the physician should do a careful clinical evaluation to determine whether the patient has thyroid dysfunction. Treatment is generally recommended when patients have either overt or subclinical hyperthyroidism because of associated adverse cardiovascular events and bone loss (8, 10). However, clinicians should be cautious about recommending life-time levothyroxine treatment when TSH is minimally raised above population-specific reference limits, particularly in elderly patients, because adverse health outcomes of minimally raised TSH concentrations, with the exception of a slow increase in risk of progression to overt hypothyroidism, have not consistently been reported (8, 10).

**Acknowledgments**

We thank Dr. Joseph G. Hollowell for providing data from NHANES III, which allowed us to develop TSH distribution curves for subpopulations of non-Hispanic Whites, non-Hispanic Blacks, and Mexican-Americans. We also thank Drs. Gilbert H. Daniels, Jayne A. Franklyn, and Eric E. Epstein for critically reading the manuscript and providing helpful suggestions.

Address all correspondence and requests for reprints to: Martin I. Surks, M.D., M.A.C.P., Montefiore Medical Center, 111 East 210th Street, Bronx, New York 10467. E-mail: msurks@westnet.com.

Disclosure Summary: The authors have nothing to disclose.

**References**

8. Surks MI, Ortiz E, Daniels GH, Sawin CT, Col NF, Cobin RH, Franklyn JA, Hershman JM, Burman KD, Denke MA, Gorman C, Cooper RS, Weissman NJ 2004 Subclinical thyroid disease: scientific review and guidelines for diagnosis and management. JAMA 291:228–238
15. Peeters RP, van Toor H, Kloorwijk W, de Rijke YB, Kuiper GG, Uitterlinden AG, Visser TJ 2003 Polymorphisms in thyroid hormone pathway genes are associated with plasma TSH and iodothyronine levels in healthy subjects. J Clin Endocrinol Metab 88:2880–2888
24. Lauberg P 1993 Persistent problems with the specificity of immunometric TSH assays. Thyroid 3:279–283
31. Hamilton TE, Davis S, Onstad L, Kopecky KJ 2008 TSH levels in a population with no clinical, autoantibody, or ultrasonographic evidence of thyroid disease: implications for the diagnosis of subclinical hypothyroidism. J Clin Endocrinol Metab 93:1224–1230
33. Boucai L, Surks MI 2009 Reference limits of serum thyrotropin (TSH) and free thyroxine (Free T4) are significantly influenced by race and age in an urban outpatient practice of medicine. Clin Endocrinol (Oxf) 70:788–793