Second-Trimester Reference Intervals for Thyroid Tests: The Role of Ethnicity

SONIA L. LA’ULU and WILLIAM L. ROBERTS

**Background:** Thyroid function changes during pregnancy, complicating the diagnosis of thyroid disorders. Maternal thyroid dysfunction has been associated with a variety of adverse outcomes. We evaluated thyroid function test results by ethnicity and week of gestation during the 2nd trimester of pregnancy.

**Methods:** We collected 3064 blood specimens in serum tubes from Asians (13%), blacks (22%), Hispanics (23%), and whites (42%). We measured thyroid-stimulating hormone (TSH), total and free thyroxine (TT4 and FT4), total and free triiodothyronine (TT3 and FT3), thyroglobulin autoantibodies (TgAb), and thyroid peroxidase autoantibodies (TPOAb) by use of an ARCHITECT i2000SR (Abbott Diagnostics). The TSH reference interval was calculated for samples negative for both TgAb and TPOAb and reference intervals for TT4, FT4, TT3, and FT3 in antibody-negative samples with normal TSH.

**Results:** Serum samples were positive for TgAb in 10.6%, 1.8%, 6.2%, 6.5%, and 5.9% of Asian, black, Hispanic, white, and combined groups, respectively. Samples were positive for TPOAb in 12.4%, 4.1%, 11.8%, 12.3%, and 10.4% of the same groups, respectively. The nonparametric reference intervals for all participants were 0.15–3.11 mIU/L (TSH), 9.3–15.2 pmol/L (0.72–1.18 ng/dL; FT4), 89.0–176.3 nmol/L (6.90–13.67 g/dL; TT4), 3.82–5.96 pmol/L (2.48–3.87 pg/mL; FT3), and 1.82–3.68 nmol/L (118–239 ng/dL; TT3).

**Conclusions:** Blacks had lower prevalences of TgAb and TPOAb positivity and of increased serum TSH. The prevalence of TgAb and TPOAb positivity was highest in Asians. Whites had the highest prevalence of increased TSH. The lower and upper reference limits of TT3 were significantly lower for Asians. Reference intervals for women in the 2nd trimester were different from those of nonpregnant individuals.

Thyroid disease is a relatively common disorder affecting women of reproductive age. Gestational hypothyroidism, which can lead to neonatal and child neurodevelopmental deficits and maternal obstetric complications, is much more common than gestational hyperthyroidism. Hormonal changes and increased metabolic demands during pregnancy produce complex alterations in thyroid hormone concentrations (1). These changes are most pronounced in the 1st trimester, but additional changes occur in both the 2nd and 3rd trimesters. For women who appear euthyroid but may have subtle thyroid dysfunction, pregnancy may result in subclinical or overt hypothyroidism. Because of the complexities of hormone changes, women should be screened for thyroid dysfunction before pregnancy or early in the 1st trimester (2). Recent findings support universal screening for thyroid dysfunction, as opposed to testing only those in high-risk groups (3).

It has been suggested that trimester-specific reference intervals be developed for thyroid function tests (4). For some analytes, particularly free thyroxine (FT4), these reference intervals likely need to be method specific (4). Ethnic background can play a role in both the prevalence of thyroid disease and the establishment of reference intervals. Thyroid-stimulating hormone (TSH) concentrations are reported to be lower in black non-Hispanics than in white non-Hispanics or Mexican Americans (5). Thyroid peroxidase autoantibodies (TPOAb) are more prevalent in whites than in blacks (5). The aims of our study were to examine the prevalence of thyroid autoantibodies.
by ethnic group in samples collected during the 2nd trimester of pregnancy and to determine reference intervals for thyroid function tests using the ARCHITECT i2000SR analyzer, by ethnic group if necessary, in those who were thyroid antibody negative.

### Materials and Methods

The following tests were performed on the ARCHITECT i2000SR (Abbott Diagnostics): TSH, FT4, total thyroxine (TT4), free triiodothyronine (FT3), total triiodothyronine (TT3), thyroglobulin autoantibodies (TgAb), TPOAb, and total β-human chorionic gonadotropin (β-hCG). All assays were performed using reagents provided by Abbott Diagnostics according to the instructions on the package insert. We used surplus specimens collected in either plain serum (red top) or serum separator tubes and submitted for 2nd trimester maternal serum testing from 396 Asians, 675 blacks, 696 Hispanics, and 1297 whites (see Table 1 in the Data Supplement that accompanies the online version of this article at http://www.clinchem.org/content/vol53/issue9). Identifying features were removed from specimens after completion of clinical testing. Before testing, specimens were thawed, mixed thoroughly, and checked for clots. The institutional review board of the University of Utah approved all studies involving human participants.

The central 95% reference intervals were determined nonparametrically for all assays using TgAb- and TPOAb-negative samples. Upper reference limits of 14.4 KU/L and 3.9 KU/L for TgAb and TPOAb, respectively, were determined previously (6). The TSH reference interval excluded TgAb- and TPOAb-positive samples. Reference intervals for FT4, TT4, FT3, and TT3 were determined using antibody-negative samples within the TSH reference interval for this group. Reference intervals were evaluated for all participants combined by week of gestation and for all weeks of gestation combined by ethnic group.

We used EP Evaluator Release 5 software (David G. Rhoads Associates) to determine reference intervals and S-PLUS software (Insightful Corp.) to calculate P values. All other statistical analyses were performed using Microsoft Excel.

### Results

TgAb- and TPOAb-positive samples were excluded from all reference interval determinations. Results of antibody prevalence by ethnicity are summarized (see Fig. 1 in the online Data Supplement). Blacks had the lowest rates of positivity for both TgAb and TPOAb. Asians had the highest rates of positivity for both TgAb and TPOAb. After exclusion of antibody-positive participants, non-parametric reference intervals were determined for TSH (Table 1). The CIs for the upper and lower reference limits for TSH for all ethnic groups combined overlapped with those for each of the individual ethnic groups.

The nonparametric reference intervals for FT4, TT4, FT3, and TT3 were determined using only antibody-negative samples with normal TSH results between 0.15 and 3.11 mIU/L. The results are summarized in Table 1. The FT4 reference interval is 9.3–15.2 pmol/L (0.72–1.18 ng/dL) for all participants combined. The CIs for the upper and lower reference limits for FT4 for all ethnic groups combined overlapped with those of each ethnic group. The TT4 reference interval is 89.0–176.3 nmol/L (6.90–13.67 μg/dL) for all participants combined. The CIs for the upper and lower reference limits for TT4 for all ethnic groups combined overlapped with those of each ethnic group. The FT3 reference interval is 3.82–5.96 pmol/L (2.48–3.87 pg/mL) for all participants combined. The CIs for the upper and lower reference limits for FT3 for all ethnic groups combined overlapped with those of each ethnic group. The TT3 reference interval is 1.82–3.68 nmol/L (118–239 ng/dL) for all participants combined and 1.60–3.31 nmol/L (104–215 ng/dL) for Asians. The CIs for TT3 reference limits for Asians did not overlap with those of the combined group.

Reference intervals were determined for all ethnic groups combined by gestational week for each analyte. The 25th, 50th, and 97.5th percentiles are plotted by analyte for gestational weeks 14–20 (Fig. 1). There were no significant trends during gestational weeks 14–20 for the median values of TSH, TT4, TT3, and FT3 (P = 0.24 to 0.47). For FT4, median values for weeks 14 and 15 were different from the other gestational weeks, but not from each other (P <0.01).

The development of hyperthyroidism requires hCG concentrations >200 000 IU/L that are sustained for several weeks (7). Of all samples analyzed, only 6 (0.2%) had total β-hCG results >200 000 IU/L. Two of those 6 samples were positive for thyroid antibodies and excluded from reference interval calculations.

Participants with increased TSH results (>3.1 mIU/L) were evaluated by ethnicity and age (see Fig. 2 in the online Data Supplement). Whites had the highest rate of increased TSH for all categories, at 5.0%, 4.0%, and 8.0% for all ages combined, <30 years of age, and ≥30 years of age, respectively. Blacks had the lowest percentage of increased TSH, showing no difference with age at 1.8%, 1.8%, and 1.9% for all ages combined, <30 years of age, and ≥30 years of age, respectively (P = 1.00). Asians also did not show a difference between increased TSH and age (P = 0.32). For Hispanics, whites, and all participants combined, a greater percentage of the population, 30 years of age or older, had increased TSH (P <0.01, P = 0.02, and P <0.01, respectively). When comparing all ages and looking for differences between ethnicities, we found a difference of increased TSH results for black vs white (P <0.01), Hispanic vs white (P = 0.04), and Asian vs black
(P = 0.01). No other differences between ethnicities were significant.

Participants with increased TSH were also evaluated by antibody status (see Table 2 in the online Data Supplement). Of all those with increased TSH, the minority (42.2%) were positive for either TPOAb or TgAb, and the majority (57.8%) were negative for both TgAb and TPOAb.

### Table 1. Summary of reference interval data.

<table>
<thead>
<tr>
<th>Ethnicity</th>
<th>n</th>
<th>2.5th percentile (lower)</th>
<th>95% CI (2.5th)</th>
<th>50th percentile (median)</th>
<th>97.5th percentile (upper)</th>
<th>95% CI (97.5th)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TSH, mIU/L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combined</td>
<td>2683</td>
<td>0.15</td>
<td>0.11–0.21</td>
<td>1.14</td>
<td>3.11</td>
<td>2.99–3.25</td>
</tr>
<tr>
<td>Asian</td>
<td>325</td>
<td>0.06</td>
<td>0.02–0.14</td>
<td>1.15</td>
<td>3.21</td>
<td>3.04–5.31</td>
</tr>
<tr>
<td>Black</td>
<td>642</td>
<td>0.07</td>
<td>0.02–0.15</td>
<td>0.97</td>
<td>2.73</td>
<td>2.49–3.04</td>
</tr>
<tr>
<td>Hispanic</td>
<td>603</td>
<td>0.19</td>
<td>0.10–0.35</td>
<td>1.20</td>
<td>3.00</td>
<td>2.77–3.40</td>
</tr>
<tr>
<td>White</td>
<td>1113</td>
<td>0.27</td>
<td>0.16–0.33</td>
<td>1.21</td>
<td>3.19</td>
<td>3.03–3.48</td>
</tr>
<tr>
<td>Adults</td>
<td></td>
<td>0.35</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FT₄, pmol/L</td>
<td></td>
<td>9.3</td>
<td>9.1–9.5</td>
<td>12.0</td>
<td>15.2</td>
<td>15.0–15.4</td>
</tr>
<tr>
<td>Combined</td>
<td>2551</td>
<td>3.82</td>
<td>3.77–3.88</td>
<td>4.85</td>
<td>5.96</td>
<td>5.88–6.05</td>
</tr>
<tr>
<td>Asian</td>
<td>296</td>
<td>3.68</td>
<td>3.47–3.77</td>
<td>4.53</td>
<td>5.59</td>
<td>5.45–6.04</td>
</tr>
<tr>
<td>Black</td>
<td>610</td>
<td>3.83</td>
<td>3.67–3.93</td>
<td>4.80</td>
<td>5.79</td>
<td>5.68–6.07</td>
</tr>
<tr>
<td>Hispanic</td>
<td>582</td>
<td>3.93</td>
<td>3.77–4.02</td>
<td>4.88</td>
<td>5.96</td>
<td>5.82–6.07</td>
</tr>
<tr>
<td>White</td>
<td>1063</td>
<td>3.87</td>
<td>3.76–3.93</td>
<td>4.94</td>
<td>6.07</td>
<td>5.96–6.18</td>
</tr>
<tr>
<td>Adults</td>
<td></td>
<td>2.63</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TT₃, nmol/L</td>
<td></td>
<td>1.82</td>
<td>1.79–1.85</td>
<td>2.62</td>
<td>3.68</td>
<td>3.62–3.74</td>
</tr>
<tr>
<td>Combined</td>
<td>2551</td>
<td>1.50</td>
<td>1.45–1.57</td>
<td>2.39</td>
<td>3.31</td>
<td>3.11–3.57</td>
</tr>
<tr>
<td>Asian</td>
<td>296</td>
<td>1.60</td>
<td>1.45–1.74</td>
<td>2.29</td>
<td>3.31</td>
<td>3.59–3.90</td>
</tr>
<tr>
<td>Black</td>
<td>610</td>
<td>1.80</td>
<td>1.72–1.88</td>
<td>2.63</td>
<td>3.77</td>
<td>3.51–3.74</td>
</tr>
<tr>
<td>Hispanic</td>
<td>582</td>
<td>1.89</td>
<td>1.72–1.96</td>
<td>2.65</td>
<td>3.59</td>
<td>3.51–3.74</td>
</tr>
<tr>
<td>White</td>
<td>1063</td>
<td>1.88</td>
<td>1.82–1.94</td>
<td>2.66</td>
<td>3.71</td>
<td>3.63–3.83</td>
</tr>
<tr>
<td>Adults</td>
<td></td>
<td>0.89</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Adult reference limits are from the ARCHITECT i2000SR package insert.

Discussion

Currently, limited information is available on suspected ethnic differences in thyroid function during pregnancy. The literature presently available has several limitations, including 1 or more of the following: sample selection not representing the US population, the entire 2nd trimester represented by 1 week or an average of gestational weeks, and/or the number of participants insufficient to calculate nonparametric reference intervals (8–14). In addition, the few studies that look at possible differences between ethnic groups also have 1 or more of the limitations mentioned above (15–17). Mandel et al. (4) described current problems associated with establishing reference intervals for pregnancy, including the responsibility of manufacturers to test specimens that represent the US population. Our study includes samples selected to represent the US population, including representation from different ethnic groups, analysis of the 2nd trimester by each gestational week individually, and adequate sample numbers (≥120 samples) for each ethnic group and gestational week to evaluate nonparametric reference intervals.

Blacks had the lowest prevalence of thyroid autoantibody positivity of all ethnic groups, with a prevalence of <5% for all categories. Asians had the highest prevalence,
with nearly 18% of the population being positive for either TgAb or TPOAb. The differences observed for antibody positivity between Asians and blacks could be explained in part by differences in age for the 2 groups. The mean maternal age for Asians was 30 years; for blacks, it was 24 years. Our results are consistent with results from other studies, which also found thyroid antibodies to be less prevalent in blacks and to increase with age (5, 12, 18–20).

There was no significant difference between ethnic groups for TSH, FT4, TT4, and FT3 reference intervals. For TT3, however, there was a statistically significant difference observed for Asians. Both the upper and lower reference limits of TT3 for Asians were lower than for all participants combined. This result may be important, because Asian women have been shown to have an increased risk of gestational thyrotoxicosis, although it is unclear what factors may account for the increased risk.
However, the prevalence of gestational thyrotoxicosis in Asian women was found to be significantly higher during weeks 8 to 11 of gestation than at 12 to 14 weeks (14). Further studies of thyroid reference intervals should be conducted during the 1st trimester of pregnancy and include Asian women.

It has been suggested that the TSH reference interval should be lower in pregnant women, particularly during the 1st trimester. An upper reference limit of 2.5 mIU/L is recommended for TSH and it is not method specific (4). We determined that the TSH upper reference limit in antibody-negative samples was 3.1 mIU/L. TT4 is predicted to increase 1.5 times during pregnancy, and 100 nmol/L (7.8 μg/dL) has been recommended as an appropriate lower cutoff for TT4 (4). Multiplying the nonpregnant reference interval by a factor of 1.5 would give a TT4 reference interval of 94–227 nmol/L (7.3–17.58 μg/dL). Our measured TT4 reference interval of 89–177 nmol/L (6.9–13.7 μg/dL) was lower, especially for the upper reference limit.

Although most changes in thyroid function occur during the 1st trimester of pregnancy, we wanted to determine whether there were any significant changes observed during the course of the 2nd trimester among all ethnic groups combined. Of note for TSH in particular, a high correlation exists between an individual’s 1st and 2nd trimester TSH measurements (21). This being a cross-sectional study, our data should provide a good representation for a diverse population during gestational weeks 14–20. When evaluating reference intervals by gestational week, analytes showed no significant change except for FT4, which at weeks 14 and 15 was different from the other gestational weeks, but not from each other. Previous studies report increases in TSH, TT3, and TT4 from the 1st trimester until late in the 2nd trimester, whereas we did not observe any significant changes across the 2nd trimester, from weeks 14 to 20 (9, 11, 17).

Another study observed significant differences for TSH, FT4, and TT3 across trimesters; however, that study used samples collected only during 1 week to represent the entire trimester (13). In agreement with our study, no significant changes in FT3 were observed after the 1st trimester (11).

High hCG concentrations can be associated with hyperemesis gravidarum, causing severe nausea and vomiting. About one-third to two-thirds of those affected with hyperemesis gravidarum have evidence of increased thyroid function, and only a small proportion of those patients have gestational thyrotoxicosis (22). Only 0.2% of our study population had total β-hCG results >200 000 IU/L. Removing the 2 samples that were antibody positive left only 0.1% having high hCG results. In response to concern regarding the potential of high hCG concentrations skewing the lower TSH reference limit, only 1 sample with a high β-hCG result had a TSH result below the lower reference limit of 0.15 mIU/L (4).

Examination of all participants with increased TSH showed whites having the highest occurrence, with 5% of the population having a TSH >3.1 mIU/L, and blacks the lowest, at 1.8%. Combining all ethnic groups resulted in increased TSH values in 3.8% of our study population. Of other studies evaluating increased TSH in pregnant populations, Walker et al. (17) reported 0.8% in American women, whereas Klein et al. (10) reported 2.5%, also in American women. Glinoer (23) found 2.2% in Belgian women, and Fukushima et al. (8) observed only 0.14% in Japanese women. Our findings best correlate with those of Quinn et al. (12), in which 3.9% of Russian women from the Samara region had increased TSH. Although the upper reference limit of TSH used for each of these studies is different, using the recommended cutoff for TSH of 2.5 mIU/L would potentially increase the rate of increased TSH results observed. This result was examined by Quinn et al. (12), who found that the percentage of participants with an abnormal TSH was originally 5.7% for the 1st trimester and increased to 31.5% when 2.5 mIU/L was used as the cutoff. Data from our study showed an increase of participants with increased TSH from 3.8% to 8.5% when the upper limit was changed from 3.1 to 2.5 mIU/L. The use of an upper reference limit of 2.5 mIU/L instead of 3.1 mIU/L for TSH will result in an additional 4.7% of women in the 2nd trimester being classified as abnormal. We think that our data support an upper reference limit of 3.1 mIU/L.

Age is not a factor for increased TSH among blacks, with 1.8% <30 years of age and 1.9% ≥30 years of age having increased TSH. Similarly, Asians also did not show an association between increased TSH and age. All other groups showed a greater percentage of those ≥30 years of age having an increased TSH. Comparison of ethnicities for increased TSH showed that there were significant ethnic differences between black vs white, Hispanic vs white, and Asian vs black.

When we examined the relationship between increased TSH and the presence of thyroid autoantibodies, nearly 60% of those with increased TSH results were negative for both TgAb and TPOAb. This finding may be explained in part by a decrease in autoantibody titers observed throughout the course of normal pregnancy (24). There was no difference in FT4 and TT4 results between those with high TSH and positive antibodies and those with high TSH and negative antibodies. Of those with increased TSH, 3.4% (4 of 116) had suppressed FT4 (<9.3 pmol/L, <0.72 ng/dL) and 3 of those 4 had positive antibodies. Considering the association of adverse outcomes to increased TSH and/or positive TPOAb, our data suggest that testing pregnant women with only 1 of these tests may be insufficient to identify women at risk for thyroid dysfunction–related pregnancy complications, particularly for ethnic groups showing a higher prevalence. Furthermore, testing should be performed as early as possible, preferably before pregnancy or during the 1st trimester (2).
All of our testing was conducted using immunoassays that are widely used in clinical laboratories. The use of tandem mass spectrometry may improve the accuracy of FT4 measurements during pregnancy (25). This technique is resistant to potential interference from heterophile antibodies, which may affect measurements by immunoassay throughout pregnancy (26, 27). Our data do not show obvious evidence of interference by heterophile antibodies, but rigorous experiments to rule out their effects were not conducted.

In conclusion, we have established 2nd trimester reference intervals for thyroid function assays on the ARCHITECT i2000sr. The prevalence of thyroid autoantibodies was lowest in blacks and highest in whites. Reference intervals for TSH, FT4, TT4, and FT3 were similar for all ethnic groups. The lower and upper reference limit of TT3 for Asians was significantly lower than all participants combined. In contrast to a previous recommendation of 2.5 mIU/L, our TSH upper reference limit is 3.1 mIU/L. Also, the current recommendation of 101 nmol/L (7.8 μg/dL) for the lower reference limit of TT4 is higher than our observed lower limit of 89 nmol/L (6.9 μg/dL). Our data show minimal change for all analytes during gestational weeks 14 to 20. The percentage of participants with increased TSH was highest for whites and lowest for blacks, whereas blacks and Asians showed no difference with increased TSH and age. However, there was a difference observed for Hispanics, whites, and combined participants, with a greater percentage of participants ≥30 years of age having an increased TSH. Also, of those with increased TSH, there was a difference between the following ethnicities: black vs white, Hispanic vs white, and Asian vs black. Of those with increased TSH, 58% were negative for both TgAb and TPOAb.

We observed differences in nonpregnant vs pregnant reference intervals for all thyroid tests, which may prove helpful in monitoring thyroid status of pregnant women affected with thyroid dysfunction. However, it may not be necessary to take into account ethnicity when evaluating thyroid status during pregnancy, because the majority of our reference intervals did not show a difference between ethnicities. Our findings, and those of others, indicate that thyroid dysfunction is more prevalent in white women, and screening may prove useful in this group before pregnancy (2, 5).

Grant/funding support: Support for this study was provided by Abbott Diagnostics and the ARUP Institute for Clinical and Experimental Pathology.

Financial disclosures: None declared.

Acknowledgments: We gratefully acknowledge Amit Phansalkar for assistance with statistical analysis and Katie Ludwig for sample collection and deidentification.

References