Narrow intra-individual variation of maternal thyroid function in pregnancy based on a longitudinal study on 132 women


1 University Department of Growth and Reproduction, GR-5064, Rigshospitalet, Blegdamsvej 9, 2100 Copenhagen, Denmark, 2 Department of Biostatistics, University of Copenhagen, DK-2400 Copenhagen, Denmark, 3 University Department of Medical Endocrinology and University Department of Clinical Biochemistry, Rigshospitalet, DK-2100 Copenhagen, Denmark, 4 Department of Obstetrics and Gynecology, Holbæk Sygehus, DK-4300 Holbæk, Denmark, and 5 University Department of Obstetrics and Gynecology, Herlev Hospital, DK-2730 Herlev, Copenhagen, Denmark

(Correspondence should be addressed to M Boas; Email: maleneboas@dadlnet.dk)

Abstract

Background: Adaptive alterations in maternal physiology cause changes in thyroid hormone levels throughout pregnancy, and precise biochemical evaluation is thus highly dependent on gestation-specific reference intervals and expected intra-individual variation.

Objective: The aim of the study was the assessment of the intra-individual variation as well as the longitudinal course of thyroid hormones during normal pregnancy and factors that influence the normal reference range for thyroid function. For this purpose, a longitudinal statistical model was applied.

Design: In a cohort of 132 pregnant women, serial blood samples were obtained and ultrasound scans were performed throughout pregnancy.

Methods: Serum levels of TSH, free and total thyroxine (T4), free and total triiodothyronine (T3) as well as autoantibodies against thyroid peroxidase and thyroglobulin were measured in 979 serum samples.

Results: Intra-individual variations of thyroid hormone concentrations were smaller than inter-individual variations (individuality index range: 0.38–0.71). Maternal height was positively associated with free T4 (FT4) (b = 0.003; P = 0.031) and pre-pregnancy body mass index with T3 and free T3 (b = 0.017; < 0.001 and b = 0.007; P < 0.001). Smoking was positively associated with T4 and FT4, but it was modulated by gestational age. Gestation-specific reference intervals for thyroid function variables from autoantibody-negative participants are presented.

Conclusions: In accordance with the data from nonpregnant adults, intra-individual variations of thyroid hormones were smaller than inter-individual variations also during pregnancy. In the evaluation of thyroid function in pregnancy, the individual longitudinal course of thyroid hormones rather than absolute values should be considered. We present a longitudinal model for the prediction of maternal thyroid function tests in pregnant women.

Introduction

Normal maternal thyroid function in pregnancy is important for fetal growth and neurological development. In the first trimester of gestation, the fetus is entirely dependent on the maternal thyroid function, which should therefore meet the needs of both fetus and mother. Throughout pregnancy, the fetus may continue to rely to some extent on maternal thyroxine (T4) supply (1), and even a marginally low T4 level in a pregnant woman may give rise to a reduction in the cognitive functions of the offspring (2–4). Hormonal changes and metabolic demands in the mother during pregnancy result in adaptive alterations in maternal thyroid function. Therefore, diagnosis as well as treatment monitoring of thyroid diseases in pregnant women depends on reliable pregnancy-specific reference intervals as well as on the knowledge about the expected change and individual variation in thyroid function tests during pregnancy. Most published reference intervals are based on cross-sectional studies without analyzing the gestation-dependent course of thyroid function tests from individual women (5–7). Therefore, we carried out a longitudinal study on 132 women throughout gestation, elucidating the intra-individual variation and factors associated with thyroid function by applying a longitudinal statistical model.

Subjects and methods

Study participants and design

From May 1999 to October 2001, all pregnant women, geographically belonging to the hospital’s primary obstetric referral area, were informed about the study...
at the first routine contact (gestational weeks 6–12) at a University Hospital in Copenhagen (Herlev Hospital). Of these, 151 women accepted to participate in the study. After informed consent, all mothers were scheduled to have three ultrasound examinations and to have blood samples obtained approximately every 4–6 weeks throughout pregnancy. Information regarding maternal diseases, smoking, parity, and parturition was obtained from medical records, one antenatal questionnaire (completed after the first trimester) and another questionnaire at birth. Thirty-seven (28%) women reported having smoked during pregnancy, whereas 90 (68%) had not smoked (5 (4%) missing answers). Twenty-nine (22%) women had stopped smoking before pregnancy and 10 (8%) during pregnancy. In the following analysis, all women who had been smoking at some time point during pregnancy were classified as smokers.

Gestational age of the newborn child was based on the sonographic measurements and the last menstrual period, as well as on the clinical evaluation of the newborn child. If there were discrepancies, sonographic measurements were used. Sixteen women dropped out of the study before term and three women were excluded: one because of serious infectious disease during pregnancy; one due to an extreme rise in TSH values during pregnancy (from 2.9 mU/l in gestational week 31 to 26.8 mU/l in gestational week 40), indicating a transition from euthyroid state to overt thyroid disease; and one due to hormonal treatment in the first trimester. Thus, the study population consisted of 132 women. In accordance with the National Academy of Clinical Biochemistry (NACB) standards (8), 28 of these women were excluded from the calculation of the reference intervals due to antibody levels above the detection limit (n = 17) or a significant disease (pregnancy-induced diabetes (n = 1), pre-eclampsia (n = 3), or hypertension (n = 7)), of which four received treatment (methyldopa, verapamil, or bendroflumethiazide). Thus, the construction of a reference interval was based on the samples from 104 women. Data regarding placental GH, insulin-like growth factor 1 (IGF1), and fetal growth in this cohort have been reported previously (9).

**Laboratory methods**

Nonfasting peripheral venous blood samples were taken from an antecubital vein between midmorning and early afternoon. Samples were centrifuged and sera were stored at −20 °C until hormone analyses.

**Assays**

TSH, T4; free (FT4), triiodothyronine (T3), and free (FT3) were measured in all the blood samples by electrochemiluminescence immunoassays on a Roche Modular E170 (Roche). Inter-assay coefficient of variations were 8.7 and 8.4% at concentrations of 0.9 and 4.9 mU/l respectively for TSH; 5.6 and 5.6% at concentrations of 81 and 167 nmol/l respectively for T4; 6.0 and 8.1% at concentrations of 12 and 30 pmol/l respectively for FT4; 6.7 and 6.6% at concentrations of 3.2 and 6.0 nmol/l respectively for T3; and 6.4 and 6.4% at concentrations of 5.3 and 15.0 pmol/l respectively for FT3. Thyroglobulin (Tg) and autoantibodies against Tg antibodies (TgAb) and thyroid peroxidase (TPOAb) were determined in the first and the last blood samples from each pregnant woman by immunofluorescent assays on the Kryptor instrument (Brahms, Hennigsdorf, Germany). The functional assay sensitivity was 25 U/ml for TgAb and was 28 U/ml for TPOAb. All samples were analyzed for TSH and thyroid hormones at the same time, and analyses were performed in 2005/2006. Placental GH and IGF1 levels were determined as described previously (9).

**Statistical analysis**

Descriptive statistics for the study population are given as medians and ranges for continuous variables and as frequencies and percentages for discrete variables. Square root transformations were applied to TSH and T4 and log transformations were applied to FT4, T3, and FT3 to approximate normal distributions. The predictors of hormone levels were analyzed in a multiple longitudinal random effects model to adjust for correlation within each subject. The nonlinear dependence of hormones on gestational age was modeled by restricted cubic splines (10). Backwards-stepwise elimination was carried out to determine the significant predictors of hormone levels. Individual deviations from the mean course of the population were approximately linear with the individual variation of intercepts and slopes. The ratios of intra-individual variation to inter-individual variation were evaluated as individuality indices ( = s.d. within women/s.d. between women). The fitted longitudinal models were validated using residual diagnostics, and were further compared to other competing longitudinal models (cubic linear splines, splines of varying orders, different transformations, and random effects exponential autocorrelation).

The reference charts were constructed by using the longitudinal random effects models including all the significant predictors found in the preceding analysis and by further adjusting for calendar time. The limits of the reference intervals were calculated as the 2.5 percentile and the 97.5 percentile of the inferred normal distribution. The longitudinal analysis was repeated for the selected population used for the calculation of reference intervals, which did not change the estimates significantly.

Statistical analyses were performed with R (11). In particular, the longitudinal models were analyzed by means of the nlme package (12).
Ethical aspects

The study was performed according to the Helsinki II Declaration and was approved by the Local Ethics Committee and the Danish Registry Agency.

Results

A total of 979 blood samples were obtained from 132 women (mean: 7.2 samples/woman; range: 1–12). All women had normal placental function as assessed by umbilical artery pulsatility index (PI) at all time points during pregnancy. The clinical characteristics of the study population are shown in Table 1. All women had singleton pregnancies and gave birth to 66 boys and 66 girls. We found no differences in thyroid function tests or clinical characteristics of the women related to maternal parity or the gender of the child. All women had one or several TSH concentrations > 2 S.D. (n = 12) had normal values of the peripheral thyroid hormones.

TgAb were measurable (above functional assay sensitivity) in 9 (7%) women and TPOAb in 15 (11%) women. Altogether, 17 (13%) women had measurable autoantibodies in at least one blood sample. Four (3.6%) women had measurable TgAb levels in the first blood sample, but not in the last blood sample, whereas one (0.9%) woman had detectable levels in only the last blood sample. For TPOAb, the numbers were 4 (3.6%) and 0 respectively. Two (1.5%) and 11 (8.3%) women had levels of TgAb or TPOAb exceeding 60 U/ml respectively, which is the level of clinical antibody positivity in our hospital. Measurable levels of autoantibodies were not associated with maternal thyroid function tests. Median Tg was 9.7 μg/l in the first half of pregnancy (gestational age < 20 weeks) and it was 12.0 μg/l in the last half of pregnancy (gestational age > 20 weeks), with three mothers (2%) having a mean Tg above 40 μg/l.

<table>
<thead>
<tr>
<th>Clinical characteristics</th>
<th>n</th>
<th>Median (range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal age (years)</td>
<td>132</td>
<td>32 (19–45)</td>
</tr>
<tr>
<td>Maternal height (cm)</td>
<td>122</td>
<td>168 (155–183)</td>
</tr>
<tr>
<td>Pre-pregnancy BMI (kg/m²)</td>
<td>120</td>
<td>22.9 (16.7–46.5)</td>
</tr>
<tr>
<td>Maternal parity</td>
<td>129</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>51</td>
<td>(40%)</td>
</tr>
<tr>
<td>≥2</td>
<td>78</td>
<td>(60%)</td>
</tr>
<tr>
<td>Maternal smoking</td>
<td>127</td>
<td></td>
</tr>
<tr>
<td>Nonsmokers</td>
<td>90</td>
<td>(71%)</td>
</tr>
<tr>
<td>Smokers</td>
<td>37</td>
<td>(29%)</td>
</tr>
<tr>
<td>Gestational age at deliveries (days)</td>
<td>132</td>
<td>281 (207–300)</td>
</tr>
<tr>
<td>Infant birth weight (kg)</td>
<td>128</td>
<td>3.56 (0.63–5.06)</td>
</tr>
<tr>
<td>Infant birth length (cm)</td>
<td>127</td>
<td>52 (41–59)</td>
</tr>
</tbody>
</table>

Discussion

All thyroid function tests showed a highly significant nonlinear dependence on gestational age (P < 0.001 for all outcomes). Maternal height was significantly positively associated with FT4 (b = 0.003; P = 0.031), whereas maternal body mass index (BMI) prior to pregnancy was positively associated with T3 and FT3 (b = 0.017; P < 0.001 and b = 0.007; P < 0.001 respectively; Fig. 1). Smoking was a significant predictor of T4 and FT4, which was modulated by an interaction with gestational age (b = 0.025; P < 0.001 and b = 0.0032; P = 0.001 respectively). Thus, the course of T4 and FT4 during pregnancy differed between smokers and nonsmokers (Fig. 1). Smokers had lower levels of T4 and FT4 than nonsmokers in the first half of pregnancy, but they had higher levels after gestational week 30. Furthermore, the date of the expected delivery was significantly negatively correlated with FT3 (b = −0.060; P < 0.001), but not with other thyroid function tests. TSH was negatively associated with total T4 (r = −0.14, P < 0.0001) and FT4 (r = −0.24, P < 0.0001).

We utilized the longitudinal design to develop a predictive model for each thyroid variable as a function of gestational age, including the significant covariates in the relevant models. When it was applied to the samples from healthy antibody-negative subjects, reference intervals were calculated. Thus, distributions and courses of thyroid function tests during pregnancy in healthy antibody-negative subjects are shown in Fig. 1A–E and Table 2. Gestation-specific S.D. of all thyroid function tests showed markedly higher inter-individual variations than the corresponding intra-individual values. Individuality indices, showing the relation between intra-individual and inter-individual variations, are shown in Table 3. Based on the longitudinal material, we estimated the expected change in FT4 from gestational weeks 10–20 and 20–30 respectively (Fig. 2).

In the questionnaires, 100 out of 125 women (76% of all women) reported having taken vitamin supplementation containing 150–200 μg iodine per day throughout most of the pregnancy. There were no significant differences in serum levels of TSH and thyroid hormones between women with and without iodine supplementation.
samples from each woman. Thus, the evaluation of the course of thyroid hormones in pregnancy as well as the associations with covariates and establishment of reference intervals was based on the factual changes of thyroid hormone concentrations in individuals rather than on the assumptions based on a cross-sectional design.

It is well known that the intra-individual variation of thyroid function tests in nonpregnant individuals is smaller than the inter-individual variation (13, 14). In pregnant women, Haddow has shown previously that TSH levels show high within-person consistency between trimesters (15). However, to our knowledge this is the first study to demonstrate that this relation applies to all thyroid function tests throughout pregnancy. The individuality index was below 0.6 for both TSH and the thyroid hormones (except for FT4, which has a slightly higher index) at all time points due to the fact that the estimated S.D.s within women in our study group were markedly lower than the estimated S.D.s between women. This indicates that reference intervals modeled for the individual would be markedly narrower than the population-based reference intervals. Consequently, population-based reference intervals are relatively insensitive in estimating an abnormality in the individual pregnant woman, either in single concentrations or in changes of hormone concentrations during gestation. In the clinical situation, it may be valuable to use the longitudinally calculated reference intervals, as they take into account the individual trends from each woman. As demonstrated in Fig. 2, the longitudinal model allows the prediction of the expected change in, as an example, FT4. Knowing the FT4 concentration at gestational week 10, the expected absolute value at week 20 or, alternatively, the expected change in FT4 between weeks 10 and 20 can be predicted. If then the actual FT4 value at week 20 differs from the prediction, the clinician is able to act upon it. This can potentially be a valuable clinical tool that can disclose a thyroid dysfunction early in its progress.

In the evaluation of anthropometric factors influencing thyroid function, we found that maternal BMI prior to pregnancy was not associated with TSH, but it was positively associated to serum levels of T3 and FT3. In contrast, previous studies have indicated positive associations between TSH and BMI in nonpregnant individuals (16–18), whereas the association between peripheral thyroid hormones and BMI remains controversial. Results from large population studies on euthyroid nonpregnant adults indicated a negative association between BMI and FT4 (16, 19), but no associations were found between BMI and FT4 (16).

Figure 1 (A) Serum TSH, (B) total T4, (C (a)) free T4, (C (b)) total T3, and (D) free T3 during pregnancy in healthy pregnant women (without thyroglobulin or thyroperoxidase antibodies). Prediction intervals are calculated considering the relevant confounders. Black lines represent medians and 2.5 and 97.5 percentiles, gray lines connect the values of each individual. BMI, pre-pregnancy body mass index.
Smoking is well known to have an impact on thyroid function in nonpregnant subjects. In our cohort, smoking was associated with changes in T4 and FT4 during pregnancy. However, these changes were neither simple nor unidirectional. Smokers tended to have lower serum concentrations of T4 and FT4 in the beginning of pregnancy, but had higher concentrations toward the end of pregnancy. In contrast, studies of nonpregnant individuals have shown higher levels of both FT4 and FT3 in smokers (20, 21) as well as lower TSH (20–23). Low levels of TSH have also been verified in smoking pregnant women (24, 25), but in accordance with our present results, several studies have found lower or similar levels of FT4 in pregnant smokers than in non-smokers (24–26). This may be due to the depressed serum levels of estrogen and human chorionic gonadotrophin (hCG) in pregnant smokers (27), which may theoretically result in attenuated hCG-mediated stimulation of the thyroid gland in the beginning of pregnancy thereby lowering thyroid hormone production. Furthermore, the lower estrogen level may result in a slower accumulation of T4 binding globulin (TBG) causing a less steep increase in total T4 in smokers.

Our cohort was recruited in an area of Denmark with a historical mild iodine insufficiency, which was corrected by the implementation of a mandatory iodine fortification program in 2000 (median urinary iodine excretion in the Copenhagen area increasing from 68 to 108 μg/l 4–5 years after the implementation of iodine fortification) (28). The main part of the blood samples in this study (90%) was collected after the initiation of the iodine fortification program, and generally the low levels of Tg indicated that the cohort as a population was regarded iodine sufficient. Median Tg in the first half of pregnancy was below 10 μg/l, which has been proposed to be a marker of iodine sufficiency (29). Furthermore, 76% of the women reported having taken individual iodine supplementation throughout pregnancy, probably ensuring the individual iodine intake. Thus, despite the early study period in relation to the national iodine fortification program, we estimate that our cohort as a population is iodine sufficient, and that our results are representative for the current iodine intake in Denmark.

Date of delivery was negatively associated with serum levels of FT3. Thus, storage time or age of the blood samples was positively associated with the concentration of FT3 in the serum samples.
This association may be due to the changes in iodine intake during the investigation period or due to the effects of freezing, thawing, and storage of blood samples, affecting FT₃ relatively more than other thyroid function tests because of the lower free concentrations of this hormone.

In the clinical evaluation of maternal thyroid function and in monitoring e.g. levothyroxine treatment of known hypothyroid pregnant women, the use of updated and precise reference intervals for thyroid function is crucial. We present reference intervals based on antibody-negative samples from a clinically well-described cohort following the guidelines from the NACB, US (8); however, we did not have the information on the family history of thyroid disease. As different covariates were associated with the levels or courses of the hormones, we have chosen to present reference curves specific for the relevant covariates in order to obtain a more accurate evaluation of thyroid function tests from an individual.

The levels of thyroid function tests in our population, especially of TSH, were generally higher than those reported previously from populations in the UK (30), Switzerland (6), the US (7), Sweden (31), and China (32) (33), but they were in accordance with previously published adult reference intervals from Denmark (34) as well as with the data from Florida, US (35), the Czech Republic (36), and India (37). These differences between reference intervals may be due to regional differences in iodine intake. However, some of the differences may depend on the differences in the statistical method as the results of the other cited studies were calculated as ± 2 S.D. of the values from the relevant trimester, whereas our results were based on longitudinal analyses. Differences from other studies may also be related to the transformation of the data or inter-laboratory differences (38, 39). This emphasizes the importance of using gestation-, assay-

<table>
<thead>
<tr>
<th>Gest. age (weeks)</th>
<th>S.D. within</th>
<th>S.D. between</th>
<th>Individuality index</th>
</tr>
</thead>
<tbody>
<tr>
<td>TSH</td>
<td>10</td>
<td>0.13</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>0.13</td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>0.13</td>
<td>0.30</td>
</tr>
<tr>
<td>T₄</td>
<td>10</td>
<td>0.45</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>0.45</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>0.45</td>
<td>1.03</td>
</tr>
<tr>
<td>Free T₄</td>
<td>10</td>
<td>0.07</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>0.07</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>0.07</td>
<td>0.11</td>
</tr>
<tr>
<td>T₃</td>
<td>10</td>
<td>0.08</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>0.08</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>0.08</td>
<td>0.16</td>
</tr>
<tr>
<td>Free T₃</td>
<td>10</td>
<td>0.07</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>0.07</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>0.07</td>
<td>0.11</td>
</tr>
</tbody>
</table>

Gest. age, gestational age.

This association may be due to the changes in iodine intake during the investigation period or due to the effects of freezing, thawing, and storage of blood samples, affecting FT₃ relatively more than other thyroid function tests because of the lower free concentrations of this hormone.

In the clinical evaluation of maternal thyroid function and in monitoring e.g. levothyroxine treatment of known hypothyroid pregnant women, the use of updated and precise reference intervals for thyroid function is crucial. We present reference intervals based on antibody-negative samples from a clinically well-described cohort following the guidelines from the NACB, US (8); however, we did not have the information on the family history of thyroid disease. As different covariates were associated with the levels or courses of the hormones, we have chosen to present reference curves specific for the relevant covariates in order to obtain a more accurate evaluation of thyroid function tests from an individual.

The levels of thyroid function tests in our population, especially of TSH, were generally higher than those reported previously from populations in the UK (30), Switzerland (6), the US (7), Sweden (31), and China (32) (33), but they were in accordance with previously published adult reference intervals from Denmark (34) as well as with the data from Florida, US (35), the Czech Republic (36), and India (37). These differences between reference intervals may be due to regional differences in iodine intake. However, some of the differences may depend on the differences in the statistical method as the results of the other cited studies were calculated as ± 2 S.D. of the values from the relevant trimester, whereas our results were based on longitudinal analyses. Differences from other studies may also be related to the transformation of the data or inter-laboratory differences (38, 39). This emphasizes the importance of using gestation-, assay-
and population-specific reference intervals in the clinical interpretation of thyroid function tests also in pregnant women.

The courses of the various thyroid function tests during pregnancy are in accordance with those reported in previous studies. TSH increased in the first trimester, which may be compatible with a prior early hCG-related suppression of TSH (40). In the remaining part of the pregnancy, TSH generally remained stable except for a slight increase near term. This final increase may be explained by the long-term effects of former mild iodine deficiency in the study area (41). Serum levels of total T₃ and total T₄ increased during the first half of gestation and then remained approximately stable at a level that was ~ 50% higher than the nonpregnant reference ranges (42). In our material, only few blood samples were obtained before gestational week 10; thus, we are unable to illustrate the expected steep increase in T₄ early in pregnancy (43). In contrast, FT₄ and FT₃ decreased until gestational week 25 and then remained stable until term. The increment of total T₁ and T₄ levels in pregnancy may be mainly due to the estrogen-induced increase in TBG, which also causes the decrease in FT₄ and T₃, stabilizing at a level that is ~ 25% lower than the nonpregnant reference levels.

In conclusion, we evaluated thyroid function longitudinally during pregnancy and established mathematical models describing the courses of the thyroid function tests. This enabled the estimations of reference intervals and intra-/inter-individual variations based on true longitudinal courses with proper weighting of samples from the participating subjects. We found that the intra-individual variation is smaller than the inter-individual variation of TSH and thyroid hormones, which is in accordance with the studies on nonpregnant individuals. For biochemical control in pregnancy, levels of the thyroid function tests should be evaluated in relation to gestation-specific reference ranges and at the same time respecting individual predicted courses, in consideration of the fact that the individual patient can be expected to maintain the same ‘track’ for each hormone in relation to the population.

Declaration of interest
There is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding
This work was supported by the European Union (Contract QLK4-CT1999-01422), the University of Copenhagen, Novo Nordisk Foundation, the Danish Medical Research Council (9700909), and the Danish Council for Strategic Research (DAN-ED. 2107-05-0006).

Acknowledgements
We are very grateful to all the pregnant women who participated in the study.

References

www.eje-online.org


23 Asvold BO, Bjoro T, Nilsen TI & Vatten LJ. Tobacco smoking and thyroid function: a population-based study. Archives of Internal Medicine 2007 167 1428–1432.


