Correlation of thyroid stimulating hormone (TSH) level with pregnancy outcome in women undergoing in vitro fertilization

Valerie L. Baker, MD, Heather M. Rone, BS, David J. Pasta, MS, H. Preston Nelson, MD, Marina Gvakharia, MD, PhD, G. David Adamson, MD

Objective: The aim of this study was to determine if pregnancy outcome for women undergoing in vitro fertilization is correlated with pre-conception thyroid-stimulating hormone level.

Study design: We performed a retrospective cohort study of in vitro fertilization cycles in our private practice with an initial positive serum human chorionic gonadotropin level and thyroid-stimulating hormone level available (n = 364). We examined whether or not birth outcome differed between cycles in which the thyroid-stimulating hormone was \( \geq 2.5 \) mIU/L compared with cycles with a thyroid-stimulating hormone level of \( \leq 2.5 \) mIU/L. Logistic regression was used to determine the association between thyroid-stimulating hormone level and spontaneous abortion rate.

Results: Delivery outcome was available for 195 cycles, 36% of which had a thyroid-stimulating hormone level \( \geq 2.5 \). The gestational age at delivery was higher in cycles with a thyroid-stimulating hormone \( \leq 2.5 \) than for cycles with a thyroid-stimulating hormone \( \geq 2.5 \) (38.5 vs 38.0 weeks for singletons, 36.0 vs 34.6 weeks for twins, overall \( P = .012 \) for thyroid-stimulating hormone level). The mean birth weight for cycles with a thyroid-stimulating hormone \( \leq 2.5 \) was higher than for cycles with a thyroid-stimulating hormone \( \geq 2.5 \) (7.33 vs 6.78 lbs for singletons, \( P = .024 \) and 5.36 vs 4.83 lbs for twins, \( P = .023 \)). Restricting analysis to cycles where the woman was not taking thyroid replacement did not change the overall conclusions. There was a trend toward increasing risk of miscarriage with increasing thyroid-stimulating hormone level in non-donor cycles, controlling for age and day 3 follicle-stimulating hormone level, but this trend did not reach statistical significance.

Conclusion: A pre-conception thyroid-stimulating hormone level \( > 2.5 \) mIU/L is associated with a lower gestational age at delivery and lower birth weight in women undergoing in vitro fertilization. © 2006 Mosby, Inc. All rights reserved.
Hypothyroidism severe enough to produce clinical symptoms has a well-documented association with both maternal and fetal complications, including increased risk of miscarriage and fetal death. More recent reports have focused on the correlation of subclinical hypothyroidism during pregnancy and adverse pregnancy outcomes including delivery before 32 weeks, placental abruption, delivery before 34 weeks, and fetal death.

Subclinical hypothyroidism has been typically defined as an asymptomatic condition associated with serum thyroid stimulating hormone (TSH) elevation in the context of normal serum free thyroxine (T4) and triiodothyronine (T3) levels. However, the normal reference range for TSH level is controversial, with some experts suggesting that “normal” ranges likely include persons with mild hypothyroidism. Given the potential adverse consequences of untreated subclinical hypothyroidism, recommendations are now being made to treat women with TSH levels that would have previously been considered normal. For example, in a publication from a recent workshop sponsored jointly by the National Center on Birth Defects and Developmental Disability (CDC) and the American Thyroid Association, it was recommended if a serum TSH > 2.5 mIU/L is confirmed before pregnancy.

Because women who have TSH levels within the reference range of ≤ 5.5 mIU/L have been considered to be normal in the laboratory that analyzes our specimens, our practice has been to not treat these women with levothyroxine before proceeding with in vitro fertilization. We therefore decided to analyze our data to determine if the pregnancy outcome for women undergoing in vitro fertilization (IVF) was associated with pre-treatment serum TSH levels.

Methods

Cycles of IVF at the Fertility Physicians of Northern California that resulted in a positive serum human chorionic gonadotropin (hCG) level and that took place between January 2003 and December 2004 and (n = 381) were considered for inclusion in the study. This time period was chosen because pre-treatment TSH levels and results for pregnancy outcome could be obtained for most cycles (n = 364). The names of all patients were deleted from the data set.

All women had an assessment of serum TSH within 1 year of beginning treatment. The study population was heterogeneous with respect to infertility diagnoses, including male factor, tubal disease, ovulation disorder with failed prior treatment, diminished ovarian reserve, endometriosis, and unexplained infertility. Many couples had multiple diagnoses. All women had a normal uterine cavity within 1 year of initiating treatment as assessed by saline hysterosonography. Prolactin levels were normal before beginning treatment for all women. Fewer than 5% of women were smokers, and all stopped smoking before beginning treatment. IVF cycles of women who had serious medical problems were not specifically excluded and comprised <5% of the total number of cycles. Egg donation cycles and frozen embryo transfer cycles were included.

Women underwent IVF after a pre-treatment cycle with oral contraceptives. Standard protocols (long luteal leuprolide acetate, gonadotropin-releasing hormone antagonist, or microdose leuprolide acetate flare) were used for ovarian stimulation. When at least two follicles reached a mean diameter of 18 mm, hCG 5000 to 10,000 IU was administered and transvaginal oocyte retrieval performed 35 hours later. Oocytes were inseminated or injected via intracytoplasmic sperm injection 4 to 5 hours after egg retrieval. Embryos were cultured in P1 media 5-10% synthetic serum substitute (Irvine Scientific, Santa Ana, CA). Intramuscular progesterone 50 mg daily was used for luteal phase support and was continued until 6 weeks of gestation. At that point, women were switched to intramuscular 17-hydroxyprogesterone 250 mg weekly until 12 weeks' gestation. Egg donor recipients were treated with a standard regimen of oral estradiol and intramuscular progesterone. Frozen embryo transfers were performed in natural cycles if a woman was ovulatory, and in controlled cycles using oral estradiol and intramuscular progesterone if a woman was anovulatory.

Serum hCG was first measured 15 days after egg retrieval. Vaginal ultrasound was performed in our clinic at 6 weeks' gestation (4 weeks after egg retrieval) and was repeated 2 weeks later. Delivery outcome was determined for each cycle by contacting the patient after the expected date of delivery. Delivery outcome was available for 195 cycles. Other pregnancies were ongoing at the time of analysis. If a patient was discharged from the clinic after her 8-week ultrasound and did not report a loss, the outcome for her cycle was classified as delivered/ongoing. For the purposes of calculating spontaneous abortion rate, biochemical and clinical miscarriages (ie, miscarriages occurring after documentation of a gestational sac) were added together.

TSH was measured in the morning using the ADVIA Centaur System (Bayer Corporation, Diagnostic Division, Tarrytown, NY) within 1 year of initiating treatment. This TSH assay is a two-site immunoassay using direct chemiluminometric technology. The reference range reported by the manufacturer was 0.35 to 5.50 mIU/L. The coefficient of variation within a run was approximately 2.5%. The run-to-run coefficient of variation was 5.31% at a mean of 0.74 mIU/L and 3.44% at a mean of 5.65 mIU/L.

Comparisons were made for birth outcome for women with TSH ≤ 2.5 mIU/L versus TSH > 2.5 mIU/L. This
In five cycles, the TSH level was not different than the mean TSH level for embryos or intracytoplasmic sperm injection (ICSI).

Results

Of the 328 women represented in 364 cycles with initial positive serum hCG and TSH level available, 295 women had one pregnancy cycle, 30 had two pregnancy cycles, and 3 had three pregnancy cycles. Baseline characteristics are presented in Table I. There were no statistically significant differences in age, gravidity, parity, number of prior spontaneous abortions, number of prior ectopic pregnancies, or baseline day 3 FSH level when comparing cycles with a TSH ≤2.5 mIU/L versus those with TSH >2.5 mIU/L (P > .15). A higher proportion of cycles with high TSH was done for women on thyroid replacement (P = .037). When the subset of cycles not using donor eggs was analyzed, there was a statistically significant difference between cycles with TSH ≤2.5 and TSH >2.5 with respect to age (35.5 vs 34.3, P = .018) but not FSH level (7.72 vs 7.79, P = .87). The difference in age would be expected to result in better outcome with higher TSH, a bias that would lead to an underestimate of the effect seen.

With respect to type of IVF performed (Table II), the difference in the percentage of cycles using egg donation was not statistically significant (P = .68). There was no difference in the proportion of cycles using frozen embryos or intracytoplasmic sperm injection (ICSI) comparing cycles with TSH ≤2.5 versus cycles with TSH >2.5 mIU/L. The mean TSH level for nondonor cycles was not different than the mean TSH level for donor cycles (1.97 vs 2.00).

The distribution of TSH results for the population is shown in the Figure. The mean TSH for the population was 1.97 ± 1.16 mIU/L (mean ± SD, median 1.71). Most women had a TSH level that was <2.0 mIU/L. In five cycles, the TSH level was >5.50 mIU/L. Four of these women had a free thyroxine level that was within the normal range, and their primary care physician recommended that they not initiate thyroid hormone or increase the dose if they were already taking thyroid hormone.

Although there was a trend for increased risk of miscarriage with increasing TSH levels, it was not statistically significant. In cycles with TSH ≤2.5 mIU/L, there were a total of 89 pregnancy losses (54 biochemical and 35 clinical losses after sonographic documentation of a gestational sac). In cycles with a TSH >2.5 mIU/L, there were a total of 31 spontaneous abortions (16 biochemical and 15 clinical). Analysis combined biochemical and clinical pregnancy losses under the category of spontaneous abortion. Donor cycles showed a lower miscarriage rate (6 of 34, 17.7%) compared with nondonor cycles (114 of 330, 34.6%; P = .046). Subsequent analysis of miscarriage was limited to nondonor cycles (to avoid heterogeneity of the data because of lower aneuploidy rate in donor cycles) and omitted the 17 cycles that ended in an ectopic pregnancy (n = 14) or therapeutic abortion for aneuploidy or anomalies (n = 3). When these nondonor cycles were examined without control for age or FSH, the relative odds ratio of miscarriage was 1.26 (95% confidence interval 0.73-2.14) with TSH >2.5 mIU/L compared with TSH ≤2.5. When age and FSH were controlled, the estimate of the odds ratio for miscarriage for TSH >2.5 compared with TSH ≤2.5 was 1.37 (0.78-2.38).

When comparing categories of TSH values 2.01 to 2.50, 2.51 to 3.0, and >3.0 with cycles with TSH ≤2.0, the estimates of the odds ratios for miscarriage compared with TSH ≤2.0 were 0.95 (0.47-1.86), 1.37 (0.60-3.05), and 1.17 (0.60-2.26), respectively. When controlling for FSH and age, the odds ratios for miscarriage with TSH levels of 2.0 to 2.50, 2.51 to 3.0 and >3.0 compared with TSH ≤2.0 were 0.92 (0.45-1.83), 1.45 (0.63-3.29), and 1.28 (0.64-2.54), respectively. Neither trend was statistically significant.

Birth outcome for women who have delivered is presented in Table III. As expected, twins delivered at an earlier gestational age and had a lower birth weight, regardless of the TSH level. The gestational age at delivery was lower for women with a TSH >2.5 mIU/L for both singletons and twins (P = .012). The effect appeared to be stronger for twins, but the difference is not statistically significant (P = .25). The birth weight for cycles with a TSH >2.5 mIU/L was lower for both singletons and twins compared with cycles with a lower TSH. A frozen cycle in a patient with a TSH of 3.22 yielded monoamniotic twins, with spontaneous abortion of one twin at 23 weeks and subsequent delivery of a singleton. Her outcome data were included in the category of live-birth/ongoing.

The percentage of cycles in which delivery occurred before 32 weeks appeared to be higher in cycles with a
TSH > 2.5 mIU/L, at least for twin pregnancies. No singleton pregnancy delivered before 32 weeks in cycles with a TSH ≤ 2.5 mIU/L compared with one singleton delivery (3%) before 32 weeks in a cycle with TSH > 2.5 mIU/L (P = .087). Among twin pregnancies, only one delivered before 32 weeks (1.8%) in cycles with a TSH > 2.5 mIU/L; but in cycles with a TSH ≤ 2.5 mIU/L, 4 of the 13 twin deliveries (31%) occurred before 32 weeks (P < .001). No patients in this study reported infection or other etiology for early delivery.

When analysis was restricted to fresh cycles for women not taking thyroid replacement, (Table IV), gestational age and birth weight for twins continued to be significantly lower in cycles with TSH > 2.5 mIU/L compared with cycles TSH ≤ 2.5 mIU/L. Statistical significance was lost for difference in gestational age and birth weight for singletons. The loss of significance for singletons in this restricted analysis was not a result of eliminating women on thyroid replacement but was instead a result of eliminating the frozen cycles and lowering the total number of cycles that were included in the analysis. For the births that resulted from frozen cycles in women not on thyroid replacement, mean gestational age for singletons in cycles with TSH > 2.5 mIU/L (n = 22) was 38.27 (32-41, 2.05) compared with a mean of 35.71 (27-39, 4.23) in cycles with TSH > 2.5 mIU/L (n = 7), P = .037. The mean birth weight for singletons from frozen cycles in women not on thyroid replacement was 7.38 lb (4.88-9.63, 1.34) with TSH > 2.5 mIU/L compared with 6.23 lb (2.06-8.19, 2.06) in cycles with TSH ≤ 2.5 mIU/L, P = .096. There was a trend for lower birth weight in twins from frozen cycles for women not on thyroid replacement that also did not reach statistical significance likely as a result of the sample size (n = 10, data not shown).

**Discussion**

The major finding of our study is that in women who become pregnant through IVF, gestational age at delivery and birth weight were lower in cycles with a TSH > 2.5 mIU/L compared with cycles with TSH ≤ 2.5 mIU/L. TSH levels > 2.5 had a possible weak positive association with spontaneous abortion rate that did not reach statistical significance. The estimated odds ratio for TSH > 2.5 mIU/L compared with TSH ≤ 2.5 mIU/L did reach 1.37 (95% confidence interval 0.78-2.38) after adjusting for age and FSH level, and this may be great enough to be of clinical significance if borne out in a study large enough to establish statistical significance.
Restricting the analysis to women not on thyroid replacement did not affect the general conclusions. What most influenced birth outcome was whether or not the TSH level was suppressed, not whether or not the woman was on thyroid replacement.

Our data are consistent with that of the recently published Camden study which found a higher relative risk of delivery before 32 weeks in women with a TSH \( \geq 3.0 \) mIU/L.\(^2\) It is important to note that the population for our study (older reproductive-age women able to afford IVF) was quite different from the Camden study (young, low-income women), and our sample size was smaller, yet the conclusions drawn were similar.

It is not certain that the effect on pregnancy outcome noted with mild elevation of TSH in our study is actually a result of low physiologic level of thyroid hormone. Other factors, such as thyroid antibodies, have been implicated in recurrent miscarriage.\(^7\) Differences in body mass index could be related to differences in TSH. The TSH level could be an epiphenomenon with some other factor being responsible for the effects on pregnancy outcome. On the other hand, maternal thyroid function may play an important role in placental implantation. T3 and epidermal growth factor act synergistically to promote trophoblast proliferation and differentiation in culture.\(^8\) Additionally, hypothyroidism leads to attenuated placental glycogen and glucose transporter metabolism. It is thus possible that maternal thyroid hormone has effects on placental function in ways that could result in smaller fetal size.

Our study has several limitations. A single TSH level for each woman was obtained. Because the variation in a single individual is likely to be minimal,\(^9\) it is unlikely that this dependence on a single value per patient significantly influenced results. None of the patients were critically ill or hospitalized, conditions that would be expected to transiently elevate the TSH. A second limitation is that only TSH values were obtained for the majority of patients. Free T4, T3, and thyroid antibody studies were not available for most cycles. Our study population was limited to women undergoing IVF, and the number of patients was modest. Because our study is retrospective, it is possible that some other factor explains the difference in outcome that has been attributed to TSH. As with any retrospective cohort study, the data collected about potential confounding variables are limited. We had minimal information about the course of prenatal care. Pregnancy outcome data are limited to that which was reported by the patient and routinely submitted to the CDC/Society for Assisted Reproductive Technology.

Our study focused on pre-conception values of TSH. TSH levels were not repeated during pregnancy, so the results cannot be directly compared with other studies in which TSH was measured during pregnancy.\(^3,4\) Although subclinical hypothyroidism rarely progresses to overt disease in nonpregnant women,\(^10\) the course of subclinical hypothyroidism during pregnancy is not well-studied. The observation that the requirement for thyroxine appears to increase for many women with hypothyroidism during pregnancy\(^11\) suggests that it is possible that women with borderline thyroid function could progress to a more serious degree of dysfunction during pregnancy. Furthermore, borderline thyroid function may also be of greater concern in women who are undergoing ovarian stimulation because it has been suggested that these women experience a drop in thyroxine and a rise in TSH during ovulation induction.\(^12,13\)

Few reports have examined an association between pre-treatment TSH levels and the outcome of IVF cycles. Cramer et al\(^14\) found that TSH levels were significantly higher in women who produced oocytes that failed to fertilize. One small study found that TSH levels were higher in women with a clinical pregnancy than in women who

<table>
<thead>
<tr>
<th>Table II</th>
<th>Type of IVF performed for the entire population considered for analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TSH ( \leq 2.5 ) mIU/L (n = 272)</td>
</tr>
<tr>
<td>Own egg</td>
<td>251 (92.3%)</td>
</tr>
<tr>
<td>Donor eggs</td>
<td>21 (7.7%)*</td>
</tr>
<tr>
<td>Fresh</td>
<td>220 (80.9%)</td>
</tr>
<tr>
<td>Frozen</td>
<td>52 (19.1%)</td>
</tr>
<tr>
<td>Insemination only</td>
<td>93 (34.2%)</td>
</tr>
<tr>
<td>ICSI only</td>
<td>138 (50.7%)</td>
</tr>
<tr>
<td>ICSI and insemination</td>
<td>41 (15.1%)</td>
</tr>
</tbody>
</table>

Data are presented as the number of cycles with percentage in parentheses. ICSI, Intracytoplasmic sperm injection; insemination, insemination of the eggs in vitro.

\* \( P = .068.\)
did not become pregnant with IVF. We did not specifically examine these questions in our study.

Screening for hypothyroidism in women planning to become pregnant or in women who are pregnant is currently controversial. The ACOG Practice Bulletin recommends that tests of thyroid function be performed in women with a personal history of thyroid disease or with symptoms of thyroid disease, not in asymptomatic women.16 In a recent workshop sponsored jointly by the National Center on Birth Defects and Developmental Disability of the CDC and the American Thyroid Association, it was recommended that treatment with levothyroxine be initiated or the dose increased if a serum TSH is confirmed before pregnancy.6 This recommendation is not evidence-based. Although no studies have been completed to confirm the benefit of this treatment, it is felt that the potential benefit-risk ratio of levothyroxine justifies its use. The authors of this statement noted that between-method biases for TSH are minimal, and they stated that there is no need for method-specific TSH ranges.

The Controlled Antenatal Thyroid Screening (CATS) study aims to enroll 22,000 women to determine if screening for thyroid function and treatment with thyroxine during pregnancy are justified. Our finding of a lower gestational age at delivery and lower birth weight in cycles with a pre-conception TSH > 2.5 mIU/L lends support to the value of pursuing this trial. Identifying women who are at risk for hypothyroidism early in pregnancy appears to be particularly important because neurological development of the offspring is critically dependent on maternal thyroid hormone during the first trimester.17,18,19

If treatment of subclinical hypothyroidism is considered based on the CATS study, care would need to be taken because concern has been raised about the potential fetal effects of overtreatment. This concern is based in part on the adverse effects of fetal exposure to high maternal hormone levels seen in mothers who have resistance to thyroid hormone.20

Because hypothyroidism is a common disorder, widespread screening could have significant public health implications. Although our study has added to emerging data suggesting a detrimental effect of mildly elevated TSH levels on pregnancy outcome, the risk versus benefit and cost-effectiveness of treating mild elevations of TSH with thyroid hormone have not been proven and warrant further study.

<table>
<thead>
<tr>
<th>Table III</th>
<th>Birth outcome for 195 cycles in which a delivery has occurred</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TSH ≤ 2.5 mIU/L (n = 150)</td>
</tr>
<tr>
<td>Singleton</td>
<td>(n = 93)</td>
</tr>
<tr>
<td>Gestational age</td>
<td>38.56 (32-41.5, 1.56)*</td>
</tr>
<tr>
<td>Birth weight</td>
<td>7.33 (4.25-9.81, 1.09)</td>
</tr>
<tr>
<td>Twin</td>
<td>(n = 57)</td>
</tr>
<tr>
<td>Gestational age</td>
<td>36.08 (30-40, 2.01)*</td>
</tr>
<tr>
<td>Birth weight</td>
<td>5.36 (2.75-7.69, 0.95)</td>
</tr>
</tbody>
</table>

Data are presented as the mean with the range and standard deviation in parentheses. The gestational age is the number of weeks at delivery. Birth weight is presented in pounds.

* P = .012 for TSH ≤ 2.5 mIU/L compared with TSH > 2.5 mIU/L.
† P = .024.
‡ P = .023.

<table>
<thead>
<tr>
<th>Table IV</th>
<th>Birth outcome for 146 cycles in which a delivery has occurred for women who transferred fresh embryos and who were not taking thyroid replacement therapy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TSH ≤ 2.5 mIU/L (n = 113)</td>
</tr>
<tr>
<td>Singleton</td>
<td>(n = 64)</td>
</tr>
<tr>
<td>Gestational age</td>
<td>38.60 (35-41, 1.36)*</td>
</tr>
<tr>
<td>Birth weight</td>
<td>7.31 (4.25-9.81, 0.99)</td>
</tr>
<tr>
<td>Twin</td>
<td>(n = 49)</td>
</tr>
<tr>
<td>Gestational age</td>
<td>35.94 (30-40, 2.09)†</td>
</tr>
<tr>
<td>Birth weight</td>
<td>5.23 (2.75-7.69, 0.93)</td>
</tr>
</tbody>
</table>

Data are presented as the mean with the range and standard deviation in parentheses. The gestational age is the number of weeks at delivery. Birth weight is presented in pounds.

* P = .89.
† P = .12.
‡ P = .032.
§ P = .025.
References

5. Surks MA, Ortiz E, Daniels GH, Sawin CT, Col NF, Cobin RH, et al. Subclinical thyroid disease: scientific review and guidelines for diagnosis and management. JAMA 2004;291:228-38.

Discussion

DENA TOWNER, MD. Normal maternal thyroid function is critical for fetal development. The best-known demonstration is of cretinism secondary to iodine deficiency. With fetal thyroid agenesis, maternal thyroxine prevents the profound developmental deficiencies as long as neonatal thyroid replacement is adequate. In women with subclinical hypothyroidism, some authors have noted a decrease in IQ points in the offspring.1 Thyroid receptors are present in the developing fetal cortex as early as 8 weeks' gestation; this suggests that thyroxine does play a role in early central nervous system development.1 As the fetal thyroid does not produce thyroid hormone that early in gestation, the source must be maternal.

Not only is maternal thyroid function necessary for neurologic development, it also has a role in placental implantation. T3 and epidermal growth factor act synergistically to promote trophoblast proliferation and differentiation.2 In addition, hypothyroidism leads to attenuated placental glucose and glucose transporter metabolism. Thus maternal thyroid hormone has an effect not only on fetal CNS development, but also on placental function in ways that could be expressed in smaller fetal size.

There is controversy over what constitutes normal TSH level: is it 5.5 mIU/L or should the upper limit be 2.5 mIU/L as recommended by some authors?3 The 5.5 mIU/L upper limit of normal is best used in an iodine-deficient population and may not be applicable to iodine-sufficient countries.

Significant thyroid hormone changes occur in the first trimester. There is a 1% to 3% increase in thyroxine daily in the first trimester. hCG production in the first trimester drives part of the increase as hCG is a weak thyroid stimulator. Each 10,000 IU of hCG leads to a decrease in TSH by 0.1 mU/mL and an increase in FT4 by 0.6 pmol/L.4 In women on replacement therapy for hypothyroidism, a recent recommendation was to increase thyroid replacement by 30% by 5 to 6 weeks' gestation.5 It could be that women with subclinical hypothyroidism simply cannot increase the production of thyroxine as needed despite the hCG effect. Research is needed in this group of women regarding pregnancy outcome, which Dr. Baker has done.