Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Casey BM, Thom EA, Peaceman AM, et al. Treatment of subclinical hypothyroidism or hypothyroxinemia in pregnancy. N Engl J Med 2017;376:815-25. DOI: 10.1056/NEJMoa1606205

Table of Contents

Investigators and Study Personnel

In addition to the authors, other members of the Eunice Kennedy Shriver National Institute of Child Health and Human Development Maternal-Fetal Medicine Units Network are as follows:

University of Texas Southwestern Medical Center, Dallas, TX – L. Moseley, J. Price, C. Melton, M. Garcia, J. Gerald, M. Santillan

University of Pittsburgh, Pittsburgh, PA – M. Cotroneo, D. DeAngelis, M. Luce, R. Kennedy, D. Nowinski

University of Alabama at Birmingham, Birmingham, AL – S. Harris, F. Biasini, M. Parks, J. Grant, C. Lee, A. Todd, K. Domnanovich, W. Andrews

Wayne State University, Detroit, MI – N. Hauff, L. Goldston, D. Driscoll

The Ohio State University, Columbus, OH – F. Johnson, S. Wylie, R. Devlin, B. Selegue, C. Latimer, J. Bauer

The University of Utah Health Sciences Center, Salt Lake City, UT – K. Hill, A. Sowles, S. Timothy, P. Reed (Intermountain Healthcare), S. Esplin (Intermountain Healthcare)

Brown University, Providence, Rhode Island – D. Allard, T. Leach, V. Watson, B. Hughes

Columbia University, New York, NY – S. Bousleiman, V. Carmona, A. Zygmunt, Y. Williams (Drexel University), M. Grant (Drexel University), C. Kitto (Christiana Care Health Systems), B. Higley (Christiana Care Health Systems), M. Falk (St. Peter's University Hospital); L. Padovano (St. Peter's University Hospital)

University of Texas Medical Branch, Galveston, TX – A. Salazar, A. Acosta, K. Smith, G. Hankins, S. Jain, M. Munn, L. Pacheco

MetroHealth Medical Center-Case Western Reserve University, Cleveland, OH – C. Milluzzi, B. Nielsen, W. Dalton, H. Cozart, E. Chien

The University of Texas Health Science Center at Houston, McGovern Medical School-Children's Memorial Hermann Hospital, Houston, TX – F. Ortiz, S. Blackwell, B. Rech, M. Hutchinson, P. Givens

University of North Carolina at Chapel Hill, Chapel Hill, NC – K. Clark, S. Timlin, K. Dorman, E. Campos, H. Byers, S. Brody (WakeMed Health & Hospitals)

Northwestern University, Chicago, IL – G. Mallett, M. Ramos-Brinson, M. Weissbourd (Lurie Children's Hospital), M. Dinsmoor (NorthShore University HealthSystem), K. Paychek (NorthShore University HealthSystem), P. Campbell

Oregon Health & Science University, Portland, OR – M. Rincon, L. Pereira, P. Blasco, S. Saxton, K. Beach, J. Snyder

George Washington University Biostatistics Center, Washington, DC – B. Jones-Binns, L. Mele.

Eunice Kennedy Shriver National Institute of Child Health and Human Development, Bethesda, MD – C. Spong, S. Tolivaisa

Thyroid Insufficiency Test Assays

Thyrotropin (TSH) assays were performed on the ADVIA Centaur (Seimens, Tarrytown, NY). The limit of detection was 0.001 uIU/mL and the interassay coefficient of variation (CV) ranged from 4.01 to 4.19 %. No significant cross-reactivity was noted with human chorionic gonadotropin levels up to 200,000 mU/mL.

Free thyroxine (fT4) assays were performed on the ADVIA Centaur (Seimens, Tarrytown, NY). The lower limit of detection was 0.1 ng/dL and the interassay CV ranged from 4.34 to 5.77 %. Cross reactivity with L-triiodothyronine (1 mg/dL), diiodotyrosine (100 mg/dL), and monoiodotyrosine (100 mg/dL) did not exceed 0.02 %.

Eligibility Criteria

Inclusion Criteria

- For the Subclinical Hypothyroidism Trial: elevated TSH (≥ 4.00 mU/L) and normal free-T_4 (0.86 to 1.90 ng/dL)
- For the Hypothyroxinemia Trial: TSH in the normal range (0.08 to 3.99 mU/L) and low free-T_4 (<0.86 ng/dL).
- Singleton pregnancy
- Gestational age at randomization no earlier than 8^0 weeks and no later than 20^6 weeks confirmed by study criteria

Exclusion Criteria

- Major fetal anomaly or demise. An ultrasound must be performed at 8 weeks 0 days or later by project gestational age and before randomization to assess fetal status.
- Planned termination of the pregnancy
- History of thyroid cancer or current thyroid disease requiring medication
- Diabetes treated with medication (insulin, glyburide).
- Collagen vascular disease (autoimmune disease) on medication, such as lupus, scleroderma and polymyalgia rheumatic.
- Receiving anticoagulant therapy.
- Depression treated with tricyclics or selective serotonin reuptake inhibitors (SSRIs).
- Other known serious maternal medical complications including:
    - Chronic hypertension requiring antihypertensive medication (including diuretics
    - Epilepsy or other seizure disorder, on medication
    - Active or chronic liver disease (acute hepatitis, chronic active hepatitis) with persistently abnormal liver enzymes
    - Cancer (including melanoma but excluding other skin cancers)

    - Heart disease (tachyrhythmia, class II or greater heart disease or on heart medication). Mitral valve prolapse without arrhythmia is not an exclusion.
- Asthma, on oral corticosteroids.
- Known illicit drug or alcohol abuse during current pregnancy.
- Planned delivery at a non-Network hospital.
- Participation in another intervention study that influences maternal and fetal morbidity and mortality, or participation in this trial in a previous pregnancy.
- Unwillingness or inability to commit to the 5 year follow-up of the infant.

Study Criteria for Determination of Gestational Age

Gestational age at trial entry was based on a comparison of gestational age according to the first day of the last menstrual period (LMP) and gestational age as assessed by the earliest dating ultrasound. If a patient had not received a dating ultrasound one had to be performed before she could be randomized. The expected date of delivery based on the study criteria could not be revised once a determination had been made.

- If the LMP date was unsure, the ultrasound measurements obtained at the patient's first dating ultrasound examination were used to determine the project gestational age, by the standard method of ultrasound gestational age determination at that institution.
- If the LMP date was sure, and the ultrasound confirmed this gestational age within the number of days specified in Table S1 'Cutoffs for Using LMP to Determine Gestational Age', then the LMP derived gestational age was used to determine the project gestational age.
- If the ultrasound determined gestational age does not confirm the LMP generated gestational age within the number of days specified in Table S1, then the ultrasound was used to determine the project gestational age.

Table S1. Cutoffs for Using LMP to Determine Gestational Age

Gestational age at first ultrasound by LMP	Ultrasound agreement with LMP
up to 19^6 weeks	$\pm$7 days
20^0 weeks or more	$\pm$14 days

Algorithm for Dose Adjustment of Thyroxine

Table S2. Dose Adjustments for the Thyroxine Group each Trial based on Monthly Testing*

Subclinical Hypothyroidism Trial		Hypothyroxinemia Trial	
TSH Level at monthly blood test (mU/L)	Thyroxine dose adjustment within 7 days (mcg/day)	Free-T_4 Level at monthly blood test (ng/dL)	Thyroxine dose adjustment within 7 days (mcg/day)
< 0.1	Decrease by 25 µg/day	< 0.86	Increase by 25 µg//day up to a maximum of 200
0.1 - 2.5 (goal)	No change	0.86 - 1.90 (goal)	No change
> 2.5	Increase by 25 µg//day up to a maximum of 200	> 1.90	Decrease by 25 µg//day

*Sham adjustments were made for the placebo group

Secondary Outcomes and Definitions

Maternal and Infant Secondary Outcomes

1. Gestational age at delivery and preterm birth < 37 weeks' gestation or < 34 weeks' gestation
2. Placental abruption: Clinical diagnosis
3. Gestational hypertension: Diastolic blood pressure ≥ 90 during pregnancy without proteinuria.
4. Preeclampsia: Diastolic blood pressure ≥ 90 during pregnancy with at least 1 + proteinuria, also includes HELLP syndrome or eclampsia
5. Gestational diabetes: Clinical diagnosis of gestational diabetes with class A1 or A2.
6. Fetal or neonatal death
7. Apgar score at 1 and 5 minutes
8. Admission to NICU (neonatal intensive care or special care nursery)
9. Small for gestational age birth weight (SGA): Birth weight less than the 10th percentile of a U.S. reference population by race/ethnicity and gender[1] (and personal communication from the author.)
10. Neonatal head circumference measured within 24 hours of birth
11. Number of days of neonatal respiratory therapy: Supplemental oxygen by endotracheal tube, nasal canula or hood; continuous positive airway pressure (CPAP); or mechanical ventilation including traditional ventilator, oscillator or jet ventilator.
12. Retinopathy of prematurity (ROP)
13. Necrotizing enterocolitis (NEC): Unequivocal presence of intramural air on abdominal x-ray, perforation seen on abdominal x-ray, clinical evidence as suggested by erythema and induration of the abdominal wall, or intra-abdominal abscess formation, or stricture formation observed at surgery or autopsy following an episode of suspected NEC. The condition is classified based on the Bell staging system.
14. Serious infectious morbidity: Clinically ill infant in whom systemic infection is suspected with a positive blood culture, CSF, or catheterized/suprapubic urine culture; or in the absence of positive cultures, clinical evidence of cardiovascular collapse or an unequivocal X-ray confirming infection.
15. Periventricular leukomalacia (PVL): Periventricular lucency in the white matter.
16. Bronchopulmonary dysplasia (BPD): Need for supplemental oxygen at 36 weeks corrected age, for babies born < 34 weeks by project gestational age only.
17. RDS: Clinical features of RDS within 24 hours of age and respiratory support and the need for supplemental oxygen therapy for six to 24 hours of age or longer.

18. Composite neonatal outcome (PVL, intraventricular hemorrhage (IVH) grade III, IV), NEC (stage II or greater), severe ROP (stage III or higher), severe RDS, BPD, neonatal death, stillbirth, and serious infectious morbidity. Severe RDS is defined as having clinical features of RDS within 24 hours of age, respiratory support and the need for supplemental oxygen therapy for six to 24 hours of age or longer, and one of the following: a full course of surfactant (2 doses of Exosurf or ≥ 3 doses of Survanta or Infasurf) or respiratory support (mechanical ventilation and/or CPAP) from 6 to 24 hours of age with a $FiO2 > 0.60$.
19. Total days in hospital nursery

[1] Alexander GR, Kogan MD, Himes JH. 1994-1996 U.S. singleton birth weight percentiles for gestational age by race, Hispanic origin, and gender. Maternal and Child Health Journal 1999 Dec;3(4):225-31.

Child Follow-up Assessments

Table S3. Follow-up Assessment Tests

Age	Assessment Instrument	Administration	Reason for Using
12 months	Bayley III	Examination of child by developmental specialist	Measures cognitive, language and motor development
24 months	Bayley III	Examination of child by developmental specialist	Measures cognitive, language and motor development
36 months	Differential Ability Scales II (DAS II)	Examination of child by developmental specialist	Measures cognitive and achievement levels of children. Correlates highly with WPPSI-R
	Child Behavior Checklist (CBCL)	Completed by parent	Assesses child's behavioral problems and social competencies as reported by parents
48 months	Connor's Rating Scales-Revised	Completed by parent	Self-report ratings to help assess ADHD and evaluate problem behavior in child
	DAS II Subtest Recognition of Pictures	Examination of child by developmental specialist	Assesses selective and sustained attention
	DAS II Subtest Recall of Digits Forward	Examination of child by developmental specialist	Assesses selective and sustained attention
60 months	Wechsler Preschool and Primary Scale of Intelligence (WPPSI-III) and subscales	Examination of child by developmental specialist	Measures IQ (primary outcome variable) and other cognitive abilities.
	Child Behavior Checklist (CBCL)	Completed by parent	Assesses child's behavioral problems and social competencies as reported by parents

The WPPSI-III is comprised of the following core subtests: Information, Vocabulary and Word Reasoning for the Verbal IQ; Block Design, Matrix Reasoning, and Picture Concepts for the Performance IQ; and Coding for Full Scale IQ. Supplementary tests include Symbol Search, Comprehension, Picture Completion, Similarities, Object Assembly; optional subtests are Receptive Vocabulary and Picture Naming. The scores derived from the WPPSI-III correlate well with the WPPSI, WISC-R, Stanford Binet (4th ed.), and McCarthy Scales.

Study Medication Adverse Events

Side effects reported by the participants are reported below. Of the 674 subclinical hypothyroidism trial participants and 520 hypothyroxinemia trial participants who had been seen for at least one study visit, 92 (13.6%) in the subclinical hypothyroidism trial and 62 (11.9%) in the hypothyroxinemia trial reported at least one side effect of the study medication. Detailed in the chart below are the most prevalent side effects reported. Side effects did not differ significantly by treatment group in either trial, P=0.28 and P=0.40 for the subclinical hypothyroidism trial and the hypothyroxinemia trial, respectively.

Table S4. Side Effects

Side Effects (%)				
	Subclinical Hypothyroidism		Isolated Hypothyroxinemia	
	Levothyroxine	Placebo	Levothyroxine	Placebo
Any side effect	15.1	12.2	10.7	13.1
Shortness of breath	0.9	1.5	0.8	0.8
Vomiting / nausea	5.3	2.4	3.8	5.8
Nervousness	1.2	0.9	0.8	0.8
Headache	3.0	2.4	1.1	1.5
Fatigue	1.2	1.2	1.5	1.9
Other	5.3	3.9*	1.5	3.9

* Includes 1 severe allergic reaction.

Additional Adverse Events

Other adverse events not described in text or tables are summarized below. Adverse events did not differ significantly by treatment group in either trial, P=0.69 and P=1.0 for the subclinical hypothyroidism trial and the hypothyroxinemia trial, respectively.

Table S5. Additional Adverse Events

	Subclinical Hypothyroidism		Isolated Hypothyroxinemia		
	Levothyroxine	Placebo	Levothyroxine	Placebo	Total
Maternal Events					
Atrial fibrillation	0	0	1	0	1
Hashimoto's thyroiditis	0	1	0	0	1
Chest/abdominal pain	1	0	0	0	1
Death at 2 years postpartum	0	0	1	0	1
Fetal Events					
Irregular fetal heart	1	0	0	0	1
IUGR, non-reassuring	0	1	0	0	1
Abnormal Doppler	0	0	1	0	1
Neonatal/Infant Events					
HIE	0	0	0	1	1
Autism Spectrum	1	0	0	0	1
Thrombocytopenia	0	0	0	1	1
Severe hearing loss	1	0	0	0	1
Total	4	2	3	2	11

Subgroup Analyses

The following are the prespecified subgroup analyses for the primary outcome.

Table S6. Tests for Interaction for the Primary Outcome

	Subclinical Hypothyroidism			Isolated Hypothyroxinemia		
Subgroups	**Levothyroxine**	**Placebo**	**P for interaction**	**Levothyroxine**	**Placebo**	**P for interaction**
TPO			0.20			0.82
TPO < 50 (IU/mL)	95.5 [91, 99]	94 [92, 96]		94 [90, 95]	91 [89, 93]	
TPO ≥ 50 (IU/mL)	98 [96, 103]	94 [91, 98]		97.5 [87, 104]	93 [85, 103]	
Gestational age at randomization			0.12			0.84
< 17 weeks	98 [95, 99]	95 [92, 97]		96 [91, 101]	92 [88, 97]	
≥ 17 weeks	95 [91, 99]	94 [92, 97]		93 [90, 95]	90 [87, 93]	
Iodine			0.76			0.37
Iodine ≤ 150 mcg/L	97 [93, 100]	94 [92, 98]		92 [87, 98]	89 [85, 92]	
Iodine ≥ 150 mcg/L	97 [93, 99]	94 [92, 97]		94 [91, 96]	93 [90, 97]	
TSH			0.81			0.98
TSH < median *	95 [92, 99]	95 [92, 97]		94 [90, 96]	92 [89, 93]	
TSH ≥ median	98 [95, 99]	94 [92, 96]		94 [90, 96]	91 [87, 96]	
			0.56			
TSH < 4.0 mU/L	97 [90, 101]	94 [91, 98]				
TSH ≥ 4.0 mU/L	97 [94, 99]	94.5 [92, 97]				
Free T4			0.48			0.63
FT4 < median **	96.5 [92, 99]	95 [92, 99]		95 [90, 96]	91 [86, 96]	
FT4 ≥ median	97 [93, 100]	94 [91, 96]		92 [89, 96]	92 [89, 95]	
Race/Ethnicity			0.45			0.48
Hispanic	90 [88, 93]	91 [88, 93]		92 [90, 95]	89 [86, 92]	
African American	95 [82, 103]	90 [83, 96]		85 [81, 90]	85 [81, 89]	
White, Other	105 [102, 109]	105 [99, 110]		103 [101, 105]	101 [98, 107]	

Data presented as median [i96% confidence interval]

*TSH medians: <4.43 mU/L for subclinical hypothyroidism group and <1.45 mU/L for hypothyroxinemia group

**FT_4 medians: <1.02 ng/dL for subclinical hypothyroidism group and <0.83 ng/dL for hypothyroxinemia group