BMJ Open

Impact of levothyroxine therapy on obstetric, neonatal and childhood outcomes in women with subclinical hypothyroidism diagnosed in pregnancy: a systematic review and meta-analysis of randomised controlled trials

Jennifer M Yamamoto,[1] Jamie L Benham,[1] Kara A Nerenberg,[2,3] Lois E Donovan[1,4]

To cite: Yamamoto JM, Benham JL, Nerenberg KA, *et al*. Impact of levothyroxine therapy on obstetric, neonatal and childhood outcomes in women with subclinical hypothyroidism diagnosed in pregnancy: a systematic review and meta-analysis of randomised controlled trials. *BMJ Open* 2018;**8**:e022837. doi:10.1136/bmjopen-2018-022837

► Prepublication history and additional material for this paper are available online. To view please visit the journal (http://dx.doi.org/10.1136/bmjopen-2018-022837).

Received 8 March 2018
Revised 9 June 2018
Accepted 11 July 2018

For numbered affiliations see end of article.

Correspondence to
Dr Lois E Donovan;
lois.donovan@albertahealthservices.ca

ABSTRACT

Objective To determine in women with subclinical hypothyroidism diagnosed in pregnancy whether levothyroxine treatment compared with control, impacts important obstetrical or childhood outcomes (specifically IQ) in randomised controlled trials.

Design Systematic review and meta-analysis.

Study eligibility criteria Randomised trials which met all the following were included: (1) reported original data of women with subclinical hypothyroidism diagnosed in pregnancy (by any prespecified study definition); (2) randomised to either levothyroxine or control (placebo or no treatment); (3) reported obstetrical outcomes and/or childhood neurodevelopmental outcomes and (4) published from 1980 to January 2018 in either English or French language.

Data sources Medline, EMBASE, CINAHL, Cochrane Database of Systematic Reviews, Cochrane Central Register of Controlled Trials and ClinicalTrials.gov.

Outcome measures Obstetrical, neonatal and childhood outcomes including: miscarriage, gestational hypertension, pre-eclampsia, preterm delivery, mode of delivery, neonatal intensive care unit admission, birth weight, gestational age at delivery, childhood IQ and neurodevelopmental scores. **Risk of bias assessment** Cochrane Risk of Bias Tool (Modified) for Quality Assessment of Randomised Controlled Trials

Results Three trials of low to unclear risk of bias with 1837 participants were included. Two studies were meta-analysed for maternal and neonatal outcomes and two studies for childhood IQ. No statistically significant differences were found for any clinical outcomes with levothyroxine therapy compared with control.

Limitations Only three trials were identified for inclusion.

Conclusions This review, based on three randomised trials in women with subclinical hypothyroidism diagnosed in pregnancy, found no evidence of benefit of levothyroxine therapy on obstetrical, neonatal, childhood IQ or neurodevelopmental outcomes. Current trial evidence does not support the treatment of subclinical hypothyroidism diagnosed in pregnancy.

Strengths and limitations of this study

- ► This systematic review provides an up-to-date summary of recently published randomised control trials and unpublished data not included in other systematic reviews and meta-analysis.
- ► A broad range of outcomes that clinicians consider during medical decision-making are analysed.
- ► This systematic review is rigorous in its design, methodology and analyses with a prepublished protocol.
- ► The inclusion of only randomised control trials minimises the bias associated with observational studies in order to clearly evaluate the effects of levothyroxine treatment in pregnant women with subclinical hypothyroidism.
- ► Only three trials were identified for inclusion.

PROSPERO registration number CRD4201707980.

INTRODUCTION

Subclinical hypothyroidism, generally defined as a thyroid stimulating hormone (TSH) value greater than the upper limit of the reference range with a normal free thyroxine (free T4) and no symptoms of hypothyroidism, is a common biochemical finding in pregnancy, occurring in more than 25% of pregnant women (depending on the assay and reference ranges used).[1] It must be noted that the diagnostic criteria for subclinical hypothyroidism in pregnancy have changed over the years and vary between countries. TSH cut-offs for subclinical hypothyroidism in pregnancy are heterogeneous, ranging from 2.5 mU/L,[2 3] to

above the 95% ile CI for the population examined and assay used, or 4 mU/L.[4]

To date, observational studies of variable methodological quality have found inconsistent associations, (positive, negative and no associations) between subclinical hypothyroidism in pregnancy and adverse obstetrical outcomes[5] including miscarriage, fetal death, preterm delivery,[6] gestational diabetes, gestational hypertension, eclampsia, placental abruption, low birth weight, 5 min Apgar score <7[7] and lower childhood IQ.[8 9] As a result, despite a paucity of high-quality randomised controlled trials (RCTs) of levothyroxine replacement therapy, many health organisations have advocated for levothyroxine replacement therapy for women with subclinical hypothyroidism diagnosed in pregnancy.[4] These recommendations are further supported by an underlying belief that levothyroxine therapy is physiologic and, thus is benign. Emerging data challenge this belief due to the demonstration of increased risks of harms associated with levothyroxine replacement and/or elevated free T4 (even in the subclinical hyperthyroid range) in pregnancy including: increased risks of pre-eclampsia,[10] small for gestational age neonates,[11] preterm delivery,[6] gestational diabetes[6] and, in fact, lower IQ.[9]

Observational studies of subclinical hypothyroidism in pregnancy have suggested two main underlying hypotheses for the association of subclinical hypothyroidism diagnosed in pregnancy and adverse pregnancy outcomes. It is important to emphasise that observational studies cannot prove causation, but rather associations. The first hypothesis is that subclinical hypothyroidism directly causes pregnancy complications through altered physiological mechanisms. Since untreated neonatal hypothyroidism is a well-established cause of impaired childhood neurodevelopment, it has been hypothesised that untreated maternal subclinical hypothyroidism may adversely impact fetal brain development. Prior to fetal production of thyroid hormone (onset at 18 weeks gestation), the only fetal source of thyroid hormone is from transplacental cross of maternal thyroid hormone.[12] Animal studies suggest the importance of thyroid hormone in fetal development of the central nervous system.[12] It is also hypothesised that subclinical hypothyroidism may impair proper uteroplacental development resulting in placental insufficiency, thereby causing adverse obstetrical outcomes. Conversely, the second hypothesis suggests that an impaired uteroplacental unit early in pregnancy directly leads to both aberrations in maternal thyroid function tests as well as pregnancy complications. In the first setting, levothyroxine therapy for subclinical hypothyroidism in pregnancy is thought to improve clinical outcomes. However, in the latter setting, levothyroxine therapy for treatment of subclinical hypothyroidism would not improve pregnancy outcomes due to a different underlying pathophysiology.

Given the lack of information about the rate of spontaneous resolution and natural history of untreated subclinical hypothyroidism in pregnancy, as well as the limitations of observational studies which do not establish causation, the best available evidence to guide clinical decision-making is RCTs of levothyroxine therapy in women with subclinical hypothyroidism diagnosed in pregnancy. As such, the objective of this review is to systematically evaluate all RCTs of levothyroxine therapy in pregnancy (compared with placebo or no treatment) to determine whether pregnant women with subclinical hypothyroidism receive benefit from levothyroxine on important obstetrical, neonatal or childhood neurodevelopmental outcomes (specifically IQ).

METHODS

We performed a systematic review and meta-analysis as outlined in the PROSPERO published protocol. Results are reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA).[13]

Study question

This review examined the following question: 'In women with subclinical hypothyroidism diagnosed in pregnancy, does treatment with levothyroxine versus control (either placebo or no treatment) influence obstetrical, neonatal or childhood outcomes?'

Study eligibility criteria

Studies that met all the following criteria were included in this review: (1) pregnant women diagnosed with subclinical hypothyroidism (by any prespecified study definition); (2) randomised to levothyroxine versus control (ie, placebo or no treatment); (3) reported obstetrical, neonatal and/or childhood neurodevelopmental outcomes; (4) reported results in either English or French and (5) published from 1980 to 25 January 2018. The publication date of 1980 was specifically chosen to coincide with the availability of more sensitive TSH assays to minimise classification bias for subclinical hypothyroidism.[14 15]

Study outcomes

The following clinically relevant obstetrical, neonatal and childhood outcomes were prespecified. Obstetrical outcomes included: miscarriage, gestational hypertension, pre-eclampsia, preterm delivery and mode of delivery. Neonatal outcomes included: admission to neonatal intensive care unit (NICU), birth weight and gestational age at delivery (weeks). Childhood outcomes included: IQ and any other validated measures of childhood neurodevelopment.

Search strategy

The search strategy was developed in consultation with a trained medical librarian. The search strategy was composed of the following terms: Hypothyroidism/or hypothyroidism.mp, Pregnancy/or pregnanc* or pregnant, Thyroxine/or thyroxine.mp or synthroid.mp or lt4.mp, and was modified according to search headings for each database. Each term was separated by the Boolean term 'AND'. The search was conducted using the Scottish Intercollegiate Guidelines Network filters for RCTs where possible.[16] The search was limited to human studies. The full search strategy is provided in online supplementary table 1.

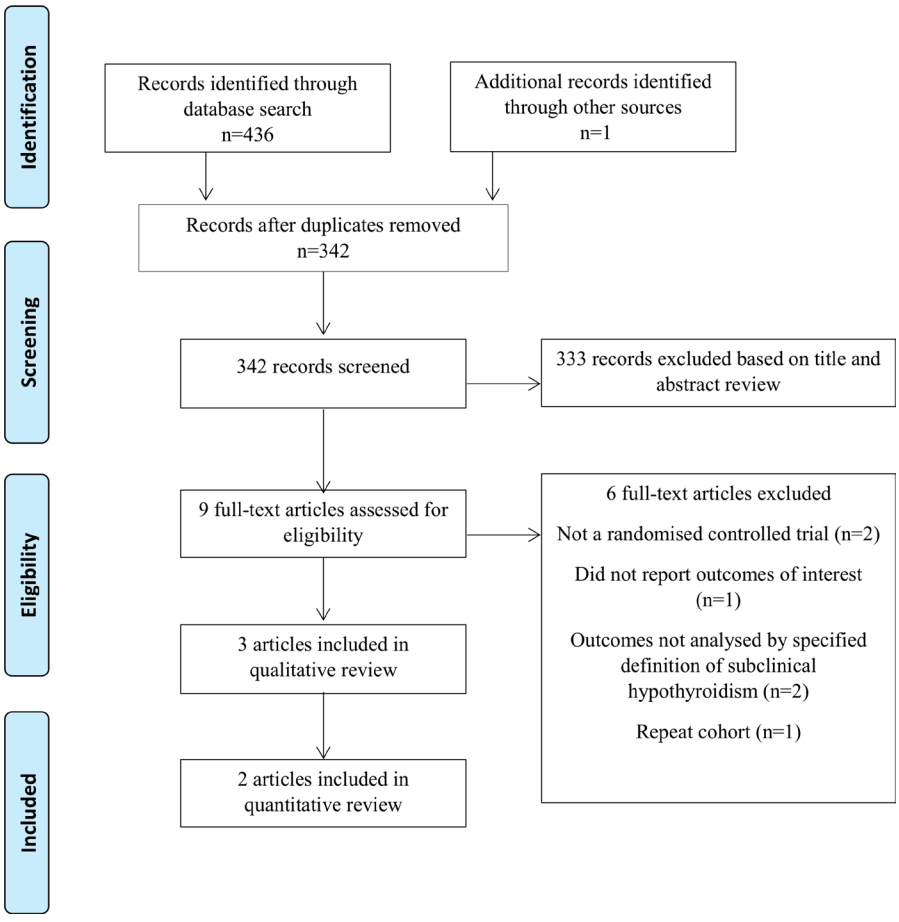


Figure 1 PRISMA flow diagram. PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

Data sources and searches

The following databases were searched in duplicate on 25 January 2018 from inception of the database: Medline, EMBASE, CINAHL, Cochrane Database of Systematic Reviews and Cochrane Central Register of Controlled Trials. In addition, references of included articles were hand-searched in duplicate to identify additional articles for inclusion. Experts in the field of endocrinology in pregnancy were consulted for potential additional articles. ClinicalTrials.gov was searched for active studies.

Study selection

Study selection was completed in two separate stages by two independent reviewers following PRISMA guidelines. In the first step, titles and abstracts of all citations retrieved by the search were reviewed for eligibility. In the second step, all citations deemed eligible at the title and abstracts stage were reviewed for eligibility in a full-text format. Reasons for exclusion were recorded. Agreement was recorded at each stage and reported as a kappa statistic. Disagreements between reviewers were resolved by consensus and by consultation with a third independent reviewer when necessary.

Data collection process

Two reviewers independently extracted relevant study information in duplicate using standardised data extraction forms. Extracted data elements included: study characteristics, study definition of subclinical hypothyroidism, baseline participant characteristics, and obstetrical, neonatal and childhood neurodevelopmental outcomes (ie, IQ or other validated neurodevelopmental outcomes). Data management was performed using Microsoft Excel.

Risk of bias assessment

Each reviewer assessed study quality independently using the Cochrane Risk of Bias Tool (Modified) for Quality Assessment of Randomised Controlled Trials.[17] This tool assesses seven methodological domains and provides a summary measure of bias for each study categorised as 'low risk of bias', 'high risk of bias' or 'unclear risk of bias'.

Data synthesis and analysis

Meta-analyses were performed using random-effect models for continuous outcomes relating to harms and benefits of treatment where there was sufficient data. Dichotomous outcomes were meta-analysed using a Mantel-Haenszel method. Risk ratios and 95% CIs were calculated. A prespecified meta-analysis of summary measures of the effect size for childhood IQ was planned. Prespecified sensitivity analyses were planned to evaluate the effects of important clinical predictors on clinical outcomes (eg, timing of initiation of levothyroxine therapy in pregnancy, dose of levothyroxine, duration of treatment, duration of follow-up, definition of subclinical hypothyroidism, treatment strategy (ie, treat to

Table 1 Characteristics of included randomised controlled trials

Study author (year)	Country	N	Intervention levothyroxine	Comparator	Definition of subclinical hypothyroidism	Percentage of trial subjects with TPO antibody ≥50 IU/mL	Clinical outcomes
Casey *et al*[18] (2017)	USA	677	Treatment of subclinical hypothyroidism or hypothyroxinaemia	Placebo	TSH >3.00 then >4.00 mU/L and a normal free T4	33%	Childhood IQ and pregnancy outcomes
Nazarpour *et al*[20] (2018)	Iran	366	Treatment of subclinical hypothyroidism	No treatment	TSH 2.5–10 mU/mL and a normal free thyroxine index	0	Pregnancy outcomes
Lazarus *et al*[19] (2012)	UK and Italy	794	Thyroid screening plus treatment if criteria met	No treatment	TSH >97.5th centile and/or free T4 <2.5th centile	Not reported	Childhood IQ

TPO, thyroid peroxidase; TSH, thyroid stimulating hormone; T4, thyroxine.

target vs fixed dose), race or ethnicity, iodine status, body mass index, method of IQ measurement, study quality and date of publication (ie, to reflect local practice)). Statistical heterogeneity was quantified using I^2 statistics where appropriate, where I^2>50% represents moderate and I^2>75% represents substantial heterogeneity across studies. Statistical analyses were performed using Review Manager (V.5.3, The Cochrane Collaboration, Copenhagen, Denmark).

Patient and public involvement

Patient involvement was not specifically sought for this study. The authors have managed a large number of women screened for subclinical hypothyroidism in pregnancy. As experienced clinicians, we recognise the burden of levothyroxine implementation includes additional laboratory testing during pregnancy and post partum and extra physician contacts as well as daily medication ingestion during pregnancy.

RESULTS

As outlined in figure 1, of the 342 citations identified by the sensitive search strategy, nine citations were identified for full-text review.[18–25] Three trials (n=1837 pregnant women) met all eligibility criteria and thus were included in the qualitative and quantitative analyses.[18–20] An additional study by Hales *et al*[26] reported a longer 9-year follow-up analysis of one of the included trials.[19] The original trial data were included in the full analysis because the reporting and follow-up in the original trial were much more complete. However, we report and discuss the 9-year follow-up data[26] because it is the longest duration of follow-up of offspring of women with subclinical hypothyroidism in pregnancy ever reported. The kappa statistic for inter-rater agreement for the title and abstract review was 0.70 (95% CI 0.45 to 0.95) and for the full-text review was 0.57 (95% CI 0.10 to 1.00).

Study characteristics

Included study characteristics are summarised in table 1. One trial compared levothyroxine to placebo[18] and two compared with no treatment.[19 20] One trial randomised women to thyroid function screening during pregnancy (followed by levothyroxine treatment if a diagnosis of subclinical hypothyroidism was found) versus thyroid function screening of stored frozen maternal blood after delivery.[19] The study definitions of subclinical hypothyroidism differed in each trial (table 1). One study included women with thyroid peroxidase (TPO) antibodies,[18] one excluded women with TPO antibodies[20] and the last study did not specify TPO antibody status.[19]

Participant characteristics

Baseline participant characteristics are summarised in table 2. Gestational age at randomisation was highest for Casey *et al*, with a mean time of randomisation of approximately 16 weeks gestation. Both Nazarpour *et al* and Lazarus *et al* randomised participants at an average gestational age closer to 12 weeks. Specifically, in the Nazarpour *et al* trial, 71% of women in the levothyroxine arm were recruited prior to 14 weeks gestation and in Lazarus *et al* 62% of women in the treatment arm were started on levothyroxine prior to 14 weeks gestation.

Risk of bias assessment

Study quality and risk of bias assessments are presented in figures 2 and 3. Only one trial[18] was considered 'low risk' in all seven domains of the Cochrane Collaboration's tool for risk of bias assessment. The Lazarus trial randomised participants to either immediate screening and treatment for subclinical hypothyroidism or delayed screening. Specifically, women randomised to immediate screening had their thyroid function assessed at approximately 12 weeks gestation. If the participants met criteria for subclinical hypothyroidism, they were started on levothyroxine. Women randomised to delayed screening had a blood sample taken at the same gestational age, but stored for analysis after delivery. Given that the results were not available until after pregnancy, no placebo was used. Thus, the Lazarus trial scored high risk for performance bias, unclear risk of detection bias, since they did not report if outcome assessment was blinded, and low risk of bias in the remaining five domains.

Maternal, obstetrical and neonatal outcomes

Data from two trials[18 20] including a total of 1043 women allowed for meta-analysis of the following five outcomes: preterm delivery <37 weeks gestation, gestational age at

Table 2 Baseline participant characteristics in included studies

Study author (year)	Maternal age (years)		Body mass index (kg/m^2)		Baseline TSH (mU/L)		Urinary iodine (µg/L)		Gestational age at trial entry (weeks)	
	Intervention arm (LT4)	Control arm	Intervention arm (LT4)	Control arm	Intervention arm (LT4)	Control arm	Intervention arm (LT4)	Control arm	Intervention arm (LT4)	Control
Casey *et al*[18] (2017)	27.7±5.7	27.3±5.7	28.1±6.4	28.2±6.4	4.5 (95% CI 4.4 to 4.7)	4.3 (95% CI 4.2 to 4.5)	199 (95% CI 184 to 238)	196 (95% CI 172 to 229)	16.6±3.0 µ	16.7±3.0 µ
Nazarpour *et al*[20] (2018)	27.0±5.34	26.9±4.74	25.8±4.95	26.0±4.64	3.8 (IQR 2.8–4.8)	3.6 (IQR 3.1–4.2)	140 (IQR 89– 219)	123 (IQR 86– 220)	11.4±4.1 α	12.2±4.3 α
Lazarus *et al*[19] (2012)	30±5.4	31±5.3	Not reported		UK: 3.8 (IQR 1.5–4.7) Turin: 3.1 (IQR 1.3–4.0)	UK: 3.2 (IQR 1.2–4.2 Turin: 2.4 (IQR 1.3– 3.9)	Not reported		12^{+3} (11^{+6}, 13^{+6}) #	12^{+3} (11^{+6}, 13^{+5}) #

Results presented as either means±SD or as median (with IQR or 95% CI) as described, LT4: levothyroxine.
Control, Placebo (Casey *et al*), no treatment (Nazarpour, Lazarus), µ=atrandomisation, α=at first visit, # at screening.
TSH, thyroid stimulating hormone.

delivery, placental abruption, NICU admission and head circumference (figure 4). There were no significant differences between levothyroxine versus control (ie, placebo or no treatment) for any of these outcomes. Meta-analysis for all outcomes had little heterogeneity (I^2=0%) except for placental abruption (I^2=73%), which demonstrated moderate heterogeneity between studies.

Additional comparisons reported in a single trial that could not be meta-analysed included stillbirth[20] (0 for both arms; p=not reported); stillbirth or miscarriage[18] (1% vs 2% for levothyroxine and placebo, respectively; p=0.36); gestational diabetes[18] (7% both arms; p=0.66); gestational hypertension[18] (10% vs 11% for levothyroxine and placebo, respectively; p=0.69); pre-eclampsia[18] (6% both groups; p=0.76) and neonatal death[18] (0 vs <1% for levothyroxine and placebo, respectively; p=0.5).

Childhood outcomes

Two trials[18 19] reported childhood IQ and other commonly used and validated measures of childhood neurodevelopment. The authors of the Casey *et al* trial provided additional unpublished data on IQ (mean and SD) allowing for inclusion in the meta-analysis (figure 5). The meta-analysis for childhood IQ showed no significant difference between levothyroxine versus control (figure 5A) (mean difference −0.69, 95% CI −2.15 to 0.78; p=0.36, I^2=0%). One trial initiated levothyroxine at a mean gestational age of 16.7 gestational weeks and reported median IQ at 5 years.[18] The other trial initiated levothyroxine at median of 13 weeks and 3 days and reported mean IQ at 3 years.[19] Both studies showed no impact of levothyroxine on Child Behaviour Checklist T-Scores (table 3). Furthermore, both trials found that childhood IQ and Child Behaviour Checklist T-Scores[18 19] were within the normal range in both levothyroxine and control arms of their trials (table 3). There were no differences in IQ in the Lazarus *et al* trial based on the components of their study definition of subclinical hypothyroidism that included women with either a TSH >97.5th centile or those with hypothyroxinaemia (free T4 <2.5th centile). Specifically, women with TSH >97.5th centile alone (n=428) had similar IQ's in their children (100.4±13.7 vs 100.7±13.0 for intervention and control arms, respectively; p=0.82) and women with a free T4 <2.5th also had similar IQ's between study arms (97.8±12.7 vs 99.2±13.3 for intervention and control arms, respectively; p=0.28).[19]

Lazarus *et al* reported no significant differences in the Behaviour Rating Inventory of Executive Function (preschool version) scale T-scores at 3 years (median 40, IQR 47–55] and 40, IQR 47–55; p=0.59 for levothyroxine and control, respectively) and normal scores in both groups.[19] Casey *et al* found no significant differences between levothyroxine and placebo arms across several additional neurodevelopmental assessment tools including: Bayley-III scores, Differential Ability Scales-II scores and Conners' Rating Scales-Revised attention deficit hyperactivity disorder score .[18]

Finally, the follow-up study of the Lazarus *et al* trial demonstrated no impact of subclinical hypothyroidism in pregnancy (using the original trial definition (ie, TSH > 97.5th

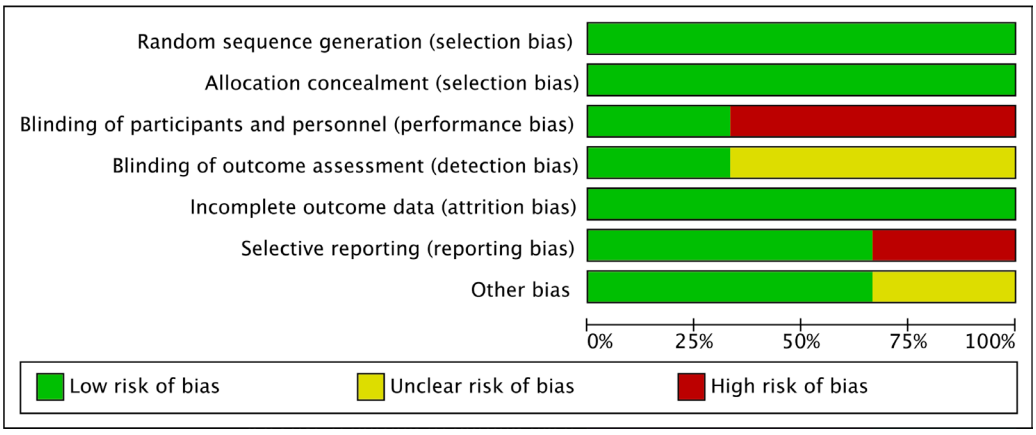


Figure 2 Risk of bias assessment.

centile and/or a free T4 <2.5th centile)) on childhood IQ[26] at 9 years of age; 101.76±12.04 (n=119) vs 102.31±13.28 (n=98) for the children of mothers treated with levothyroxine and untreated, respectively. There was also no IQ difference after 9 years of follow-up among children of mothers who only had TSH >97.5th centile (103.06±12.20 (n=63) vs 104.80±12.94 (n=51) levothyroxine and untreated, respectively). When a sensitivity analysis was done by substituting the 9-year follow-up data for the 3-year follow-up data from this same cohort the result remained unchanged and showed no impact of levothyroxine on childhood IQ (mean difference −0.5, 95% CI −2.46 to 1.46; p=0.62, I^2=0%) (figure 5B).

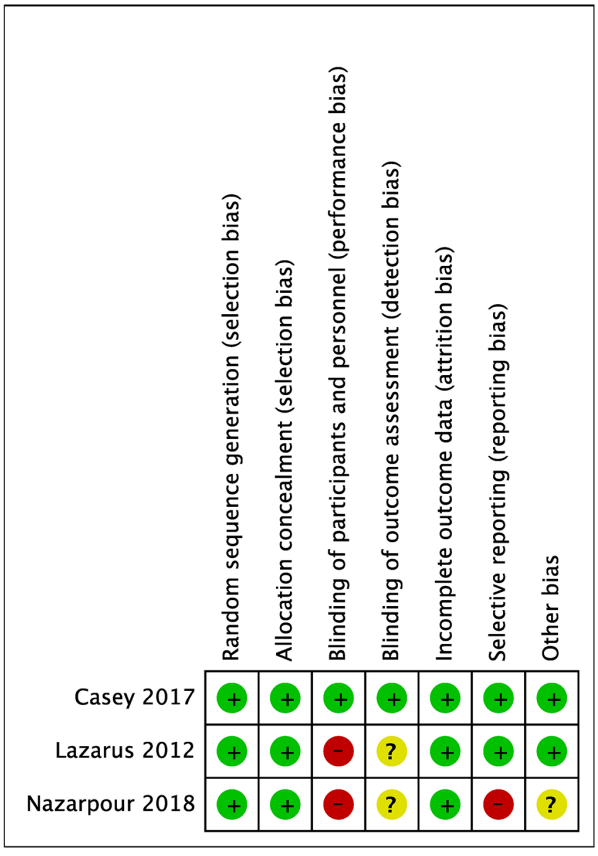


Figure 3 Trial quality assessment.

Potential harms associated with treatment of subclinical hypothyroidism

Lazarus *et al* reported that 10% of women in their levothyroxine arm had their dose of levothyroxine lowered due to either biochemical hyperthyroidism (ie, a very low TSH or a high free T4 level) or clinical symptoms of hyperthyroidism (ie, palpitations) associated with levothyroxine overtreatment.[19] The other two trials did not report similar data. The Casey *et al* trial reported similar rates of anticipated adverse events between treatment arms; 15.1% vs 12.2% in the levothyroxine and placebo arms, respectively.[18] Unanticipated adverse events in the levothyroxine arm were: maternal chest/abdominal pain (n=1), irregular fetal heart (n=1), autism spectrum (n=1) and severe hearing loss (n=1) whereas the placebo arm reported Hashimoto's thyroiditis (n=1) and intrauterine growth restriction (n=1).[18] Nazarpour *et al*[20] did not report adverse events.

Sensitivity analysis

Our prespecified sensitivity analysis could not be performed because of the small number of studies reporting the specific data. One trial evaluated the TSH cut-off of ≥4 mU/L vs <4 mU/L in a post hoc analysis.[20] Among women with a TSH ≥4 mU/L, there was a lower proportion of preterm delivery in the levothyroxine treated group compared with the untreated group (7.3% vs 19%, respectively; p value not reported). Whereas in women with a TSH <4 mU/L, levothyroxine treatment was associated with a higher preterm delivery rate compared with untreated women (12.8% vs 8.3%, respectively; p value not reported). We could not perform a sensitivity analysis as the number of women who had a TSH ≥4 mU/L was not reported in this post hoc analysis. Only three studies met our inclusion criteria, which all had negative results relating to the primary outcome, therefore, a formal assessment of publication bias was not done.

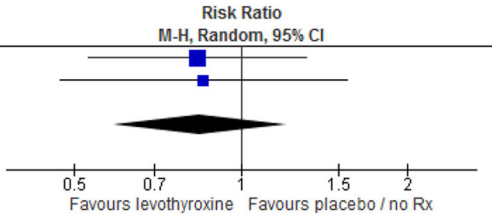
DISCUSSION

This systematic review and meta-analysis of RCTs of levothyroxine treatment in women with subclinical hypothyroidism diagnosed in pregnancy demonstrated no evidence of benefit of levothyroxine treatment across a wide range of

A.

Study or Subgroup	Levothyroxine Events	Total	Placebo or no treatment Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Casey 2017	31	339	37	338	63.3%	0.84 [0.53, 1.31]
Nazarpour 2018	18	183	21	183	36.7%	0.86 [0.47, 1.55]
Total (95% CI)		**522**		**521**	**100.0%**	**0.84 [0.59, 1.21]**
Total events	49		58			

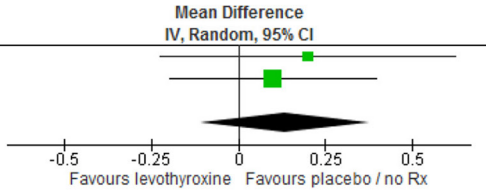
Heterogeneity: Tau² = 0.00; Chi² = 0.00, df = 1 (P = 0.95); I² = 0%
Test for overall effect: Z = 0.93 (P = 0.35)



B.

Study or Subgroup	Levothyroxine Mean	SD	Total	Placebo or no treatment Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Casey 2017	39.1	2.5	339	38.9	3.1	338	32.9%	0.20 [-0.22, 0.62]
Nazarpour 2018	38	1.4	183	37.9	1.5	183	67.1%	0.10 [-0.20, 0.40]
Total (95% CI)			**522**			**521**	**100.0%**	**0.13 [-0.11, 0.38]**

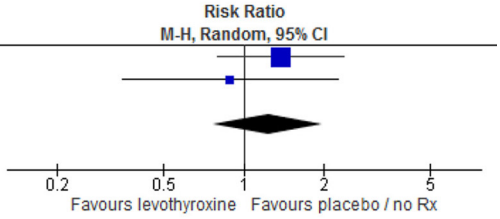
Heterogeneity: Tau² = 0.00; Chi² = 0.14, df = 1 (P = 0.71); I² = 0%
Test for overall effect: Z = 1.07 (P = 0.28)



C.

Study or Subgroup	Levothyroxine Events	Total	Placebo / no treatment Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Casey 2017	29	339	21	338	74.7%	1.38 [0.80, 2.37]
Nazarpour 2018	8	183	9	183	25.3%	0.89 [0.35, 2.25]
Total (95% CI)		**522**		**521**	**100.0%**	**1.23 [0.77, 1.97]**
Total events	37		30			

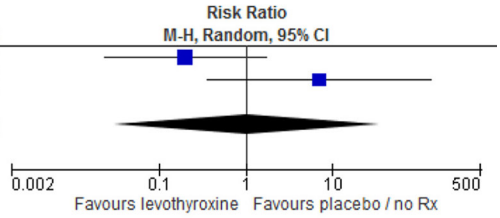
Heterogeneity: Tau² = 0.00; Chi² = 0.64, df = 1 (P = 0.43); I² = 0%
Test for overall effect: Z = 0.88 (P = 0.38)



D.

Study or Subgroup	Levothyroxine Events	Total	Placebo or no treatment Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Casey 2017	1	339	5	338	54.2%	0.20 [0.02, 1.70]
Nazarpour 2018	3	183	0	183	45.8%	7.00 [0.36, 134.56]
Total (95% CI)		**522**		**521**	**100.0%**	**1.02 [0.03, 33.52]**
Total events	4		5			

Heterogeneity: Tau² = 4.67; Chi² = 3.69, df = 1 (P = 0.05); I² = 73%
Test for overall effect: Z = 0.01 (P = 0.99)



E.

Study or Subgroup	levothyroxine Mean	SD	Total	placebo or no treatment Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Casey 2017	33.8	1.8	339	33.9	1.7	338	57.7%	-0.10 [-0.36, 0.16]
Nazarpour 2018	34.6	1.4	183	34.7	1.6	183	42.3%	-0.10 [-0.41, 0.21]
Total (95% CI)			**522**			**521**	**100.0%**	**-0.10 [-0.30, 0.10]**

Heterogeneity: Tau² = 0.00; Chi² = 0.00, df = 1 (P = 1.00); I² = 0%
Test for overall effect: Z = 0.98 (P = 0.33)

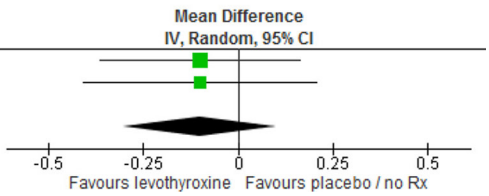


Figure 4 Results of meta-analysis of effects of treatment with levothyroxine on clinical outcomes, (A) preterm delivery <37 gestational weeks, (B) gestational age at delivery, (C) neonatal intensive care unit admission, (D) placental abruption, (E) head circumference.

prespecified maternal, obstetrical and neonatal outcomes and childhood IQ. Importantly, this review found no evidence of benefit of levothyroxine treatment during pregnancy on childhood IQ and neurodevelopment measures at 3, 5 or 9 years of follow-up. These findings challenge the current clinical practice guidelines which recommend the use of levothyroxine therapy for treatment of women with subclinical hypothyroidism diagnosed in pregnancy.[4] As outlined below, these results also question the recommendations of levothyroxine therapy earlier in pregnancy than occurred in the available RCTs and at lower TSH thresholds among women who are TPO antibody positive.[4 27 28]

This review found that the timing of initiation of levothyroxine therapy was not significantly associated with

Table 3 Neurodevelopmental outcomes of studies

Study (year) [Childhood age]	Childhood IQ		Child Behaviour Checklist T-Score at 3 years ^		Child Behaviour Checklist at 5 years	
	Levothyroxine	Control	Levothyroxine	Control	Levothyroxine	Control
Casey *et al*[18] (2017) [5 years]	n=311 95.9±15.7[#]	n=314 96.4±14.8[#]	n=306 46 (95% CI 45 to 48)	n=309 46 (95% CI 45 to 48)	n=314 44 (95% CI 43 to 46)	n=313 44 (95% CI 42 to 46)
Lazarus *et al*[19] (2012) [3 years]	n=390 99.2±13.3	n=404 100.0±13.3	n=390 44.4±12.4	n=404 45.1±13.6		
Hales *et al*[26] (2018) [9 years]	n=119 101.8±12.0	n=98 102.3±13.3				

Results presented as either means±SD or as median (with IQR or 95% CI) as described. [#] unpublished data provided from the Casey et al trial, ^ score of < 60 was defined as normal range.

childhood IQ. First, in the Lazarus *et al* trial,[19] 62% of women were started on levothyroxine at less than 14 weeks gestation. There was no difference in childhood IQ among women who were started on levothyroxine prior to 14 weeks gestation compared with no treatment (98.7±14.1 vs 100.00±13.3 levothyroxine vs control, respectively; p=0.28), although this study was underpowered to examine this difference.[19] Second, a prespecified subgroup analysis in the Casey *et al* trial found no significant interaction for childhood IQ by gestational weeks at randomisation[18] but likely was also underpowered to answer this question. Third, the Nazarpour *et al* trial entry visit was at the earliest gestation, a mean of 11.4 (SD ±4.1) weeks. This trial also showed no impact of early timing of initiation of levothyroxine therapy. Given the above findings, it appears implausible that treatment at an earlier gestational age may improve childhood IQ, particularly given the finding of normal childhood IQ up to 9 years later.[18 19 26] Furthermore, childhood IQ was not different between children of mothers with normal thyroid function and those with subclinical hypothyroidism after 9 years of follow up.[26]

Some have hypothesised that levothyroxine treatment prepregnancy might be more appropriate since that might affect placentation and very early fetal neurodevelopment. The T4life[29] and TABLET (ThyroidAntiBodies and LEvoThyroxine) trials,[30] results of which are yet to be reported, are assessing the impact of prepregnancy levothyroxine on pregnancy outcomes in TPO antibody positive women. However, these trials include euthyroid women only. Interestingly, a recent large randomised control trial of preconception levothyroxine in TPO antibody positive women with TSH levels between 0.5 and 4.78 mIU/L undergoing in vitro fertilisation showed no benefits of levothyroxine on pregnancy, miscarriage, live birth or preterm delivery rates[31] suggesting no benefit of levothyroxine replacement on placental development.

Based on their observational studies, some investigators[27 32] have raised the possibility of a synergistic association of subclinical hypothyroidism and TPO antibody positivity to adversely influence obstetrical and childhood outcomes. The trials in our review that included thyroid antibody positive women[18 19] were not specifically powered to analyse the results by maternal antibody status. The Casey study, however, had a large proportion of participants with TPO antibody positivity (33% of women with subclinical hypothyroidism)[33] and found no association between TPO antibodies and childhood IQ in either study arm.[18] Furthermore, in addition to their published trial results, Casey *et al*'s published

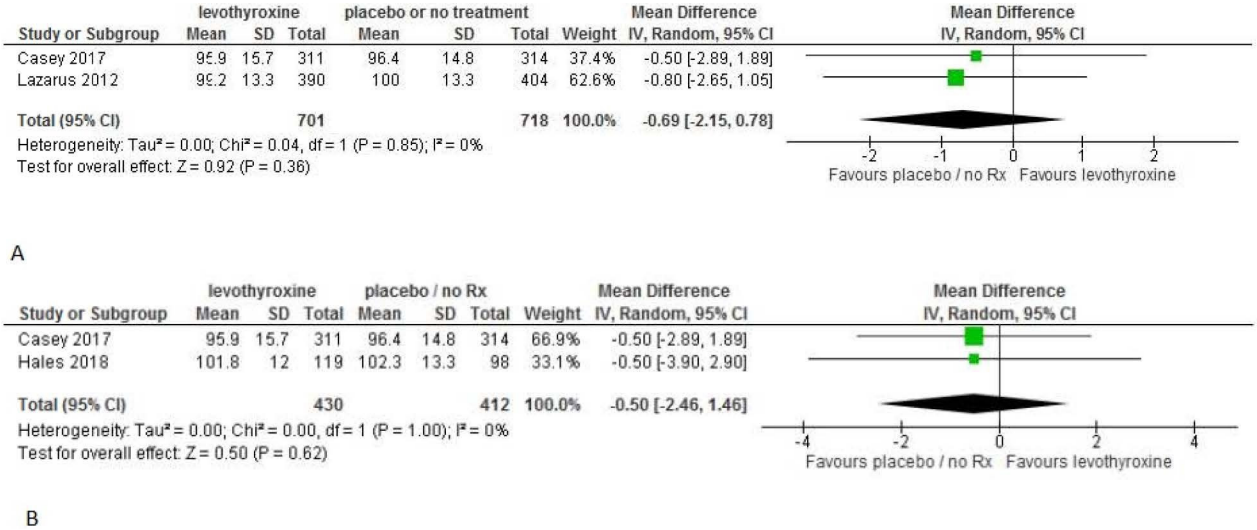


Figure 5 Results of meta-analysis of effects of treatment with levothyroxine on childhood IQ (A) childhood IQ at 3 and 5 years, (B) sensitivity analysis of childhood IQ at 5 and 9 years (note all three studies were not combined because the Lazarus and Hales trials were from the same cohort).

abstract of their trial demonstrated no impact of TPO status on pregnancy loss or stillbirth, preterm birth, pre-eclampsia or placental abruption.[33] The Lazarus *et al* trial[19] did not specifically report thyroid antibody status, nor was it part of their trial exclusion criteria. Based on other studies of reproductive age women, it is reasonable to assume that at least 14% of women in this trial would have had positive antithyroid antibodies.[18 34] Thus, trial evidence to date shows no benefit of treatment with levothyroxine for women with subclinical hypothyroidism who are TPO antibody positive during pregnancy.

The included RCTs did not consistently report harms of levothyroxine treatment and were also of inadequate sample size to fully examine this association. This is important as large observational studies of levothyroxine therapy in pregnancy[6] as well as high free T4 values in pregnancy[10] have found increased risks of gestational diabetes, pre-eclampsia, and small for gestational age[6 10 11] Only the Casey trial[18] (n=677) reported these outcomes and found no increase in gestational diabetes, pre-eclampsia and small for gestational age in the levothyroxine arm.[18] The only direct harm of levothyroxine found by this review was iatrogenic hyperthyroidism in 10% of the treated women (either biochemical or symptomatic) in the Lazarus trial,[19] which used higher dosing of levothyroxine compared with the other trials. Reassuringly, the 9-year follow-up analysis suggests that this oversupplementation with levothyroxine did not have a detrimental effect on childhood IQ[26].

None of the included studies examined the cost-effectiveness of treatment with levothyroxine for subclinical hypothyroidism in pregnancy. Given the lack of evidence of benefit of levothyroxine on obstetrical, neonatal and neurodevelopmental outcomes, the costs of healthcare resources associated with screening, therapy and monitoring of subclinical hypothyroidism in pregnancy and post partum must be rigorously evaluated to ensure appropriate healthcare resource allocation.

This systematic review has several strengths. It provides an up-to-date summary of recently published RCTs not included[20 26] in other systematic reviews and meta-analyses[5 35 36] and it addresses a broad range of clinical outcomes that clinicians consider during medical decision-making. This systematic review is rigorous in its design, methodology and analyses. Furthermore, the inclusion of only RCTs minimises the bias associated with observational studies[35 36] in order to clearly evaluate the effects of levothyroxine treatment in pregnant women with subclinical hypothyroidism.

Our review's findings must be interpreted in the context of the study's limitations. First, this review limited the search to trials published in English or French. This is unlikely to have biased the results as the included trials were conducted across several countries (North America, Europe and the Middle East) and included mixed racial populations with differing maternal iodine status which also did not impact results (table 1). Thus, our study's results are likely generalisable to a broad range of populations. Second, our review limited study selection to published trials (a priori). These eligibility criteria were specifically chosen to ensure high-quality results, though limits the ability to use data from the Casey and Lazarus trials[18 19] that are currently published only in abstract form.[24 33] Of note, the Lazarus abstract also reports no significant differences in obstetrical outcomes among women with subclinical hypothyroidism randomised to levothyroxine compared with untreated women.[24] Similarly, analysis of maternal and neonatal outcomes by maternal TPO antibody status in the Casey trial is currently published in abstract form[33] and also showed no association of TPO antibody status on maternal outcomes. Once these results are published in full-text format, an updated systematic review will be important. A priori we decided to use random-effects models. Only two studies per outcome were available for meta-analyses so between study variance could not be accounted for accurately. When fixed-effect models (with inverse weighting) were used, the results were the same as those from the random-effects models. Lastly, given the median and IQR or 95% CI of baseline TSH measurements for women enrolled in these trials, it remains difficult to extrapolate these trial findings to women with subclinical hypothyroidism in pregnancy with TSH measurement between 5 and 10 mU/L. Therefore, a trial of women with subclinical hypothyroidism with TSH >5.0 to <10.0 mU/L may be warranted.

In summary, data from three randomised trials fail to demonstrate any evidence of benefit of treating subclinical hypothyroidism in pregnancy on important obstetrical, neonatal, childhood IQ and neurobehavioural outcomes. Current trial evidence does not support the treatment of subclinical hypothyroidism diagnosed in pregnancy. Given the lack of evidence of benefit of levothyroxine therapy for subclinical hypothyroidism in pregnancy, the treatment and by extension the practice of screening for subclinical hypothyroidism in pregnancy needs careful re-examination.

Author affiliations
[1]Division of Endocrinology and Metabolism, Department of Medicine, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada
[2]Division of General Internal Medicine, Department of Obstetrics and Gynaecology, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada
[3]Division of General Internal Medicine, Department of Medicine, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada
[4]Division of Endocrinology and Metabolism, Department of Obstetrics and Gynaecology, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada

Acknowledgements We thank Helen Lee Robertson (University of Calgary, Calgary, Canada) for her expert advice on our search strategy. We also thank Lonnie Zwaigenbaum, Professor and Director of Autism Research in the Department of Pediatrics at the University of Alberta for sharing his expertise on the interpretation of childhood IQ and neurodevelopmental measures. We also thank the authors of the Casey *et al* trial for providing additional unpublished data on IQ (mean and SD).

Contributors JMY and LED are endocrinologists and clinical researchers who conceived the idea for this systematic review and meta-analysis. KAN, a specialist in obstetric medicine, contributed content and methodological expertise to this review. JLB contributed her expertise in systematic review methodology to the design and conduct of the systematic review. All authors participated in the study design. JMY and JLB conducted the literature searches, study selection and data extraction. LED and JMY contributed to writing the first draft of the manuscript. All authors contributed to critical review and approval of the final manuscript. LED is the guarantor of this article.

Funding This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement This systematic review and meta-analysis used data from published trials referenced in this article. At our request the National Institute of Child Health and Human Development, Maternal Fetal Medicine Unit Network provided unpublished data on childhood IQ (mean and SD) from the Casey *et al* trial.

REFERENCES

1. Li C, Shan Z, Mao J, *et al*. Assessment of thyroid function during first-trimester pregnancy: what is the rational upper limit of serum TSH during the first trimester in Chinese pregnant women? *J Clin Endocrinol Metab* 2014;99:73–9.
2. De Groot L, Abalovich M, Alexander EK, *et al*. Management of thyroid dysfunction during pregnancy and postpartum: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab* 2012;97:2543–65.
3. Lazarus J, Brown RS, Daumerie C, *et al*. 2014 European thyroid association guidelines for the management of subclinical hypothyroidism in pregnancy and in children. *Eur Thyroid J* 2014;3:76–94.
4. Alexander EK, Pearce EN, Brent GA, *et al*. 2017 Guidelines of the American Thyroid Association for the Diagnosis and Management of Thyroid Disease During Pregnancy and the Postpartum. *Thyroid* 2017;27:315–89.
5. Negro R, Stagnaro-Green A. Diagnosis and management of subclinical hypothyroidism in pregnancy. *BMJ* 2014;349:g4929.
6. Maraka S, Mwangi R, McCoy RG, *et al*. Thyroid hormone treatment among pregnant women with subclinical hypothyroidism: US national assessment. *BMJ* 2017;356:i6865.
7. Maraka S, Singh Ospina NM, O'Keeffe DT, *et al*. Effects of Levothyroxine Therapy on Pregnancy Outcomes in Women with Subclinical Hypothyroidism. *Thyroid* 2016;26:980–6.
8. Haddow JE, Palomaki GE, Allan WC, *et al*. Maternal thyroid deficiency during pregnancy and subsequent neuropsychological development of the child. *N Engl J Med* 1999;341:549–55.
9. Korevaar TI, Muetzel R, Medici M, *et al*. Association of maternal thyroid function during early pregnancy with offspring IQ and brain morphology in childhood: a population-based prospective cohort study. *Lancet Diabetes Endocrinol* 2016;4:35–43.
10. Medici M, Korevaar TI, Schalekamp-Timmermans S, *et al*. Maternal early-pregnancy thyroid function is associated with subsequent hypertensive disorders of pregnancy: the generation R study. *J Clin Endocrinol Metab* 2014;99:E2591–E2598.
11. León G, Murcia M, Rebagliato M, *et al*. Maternal thyroid dysfunction during gestation, preterm delivery, and birthweight. The Infancia y Medio Ambiente Cohort, Spain. *Paediatr Perinat Epidemiol* 2015;29:113–22.
12. de Escobar GM, Obregón MJ, del Rey FE. Maternal thyroid hormones early in pregnancy and fetal brain development. *Best Pract Res Clin Endocrinol Metab* 2004;18:225–48.
13. Moher D, Liberati A, Tetzlaff J, *et al*. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009;1519:264.
14. Weeks I, Sturgess M, Siddle K, *et al*. A high sensitivity immunochemiluminometric assay for human thyrotrophin. *Clin Endocrinol* 1984;20:489–95.
15. Nicoloff JT, Spencer CA. Clinical review 12: The use and misuse of the sensitive thyrotropin assays. *J Clin Endocrinol Metab* 1990;71:553–8.
16. Scottish Intercollegiate Guidelines Network. Randomised Controlled Trials. 2017 http://www.sign.ac.uk/search-filters.html (accessed 19 Oct 2017).
17. The Cochrane Collaboration. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0, 2011. http://handbook.cochrane.org
18. Casey BM, Thom EA, Peaceman AM, *et al*. Treatment of Subclinical Hypothyroidism or Hypothyroxinemia in Pregnancy. *N Engl J Med* 2017;376:815–25.
19. Lazarus JH, Bestwick JP, Channon S, *et al*. Antenatal thyroid screening and childhood cognitive function. *N Engl J Med* 2012;366:493–501.
20. Nazarpour S, Ramezani Tehrani F, Simbar M, *et al*. Effects of Levothyroxine on Pregnant Women With Subclinical Hypothyroidism, Negative for Thyroid Peroxidase Antibodies. *J Clin Endocrinol Metab* 2018;103:926–35.
21. Ma L, Qi H, Chai X, *et al*. The effects of screening and intervention of subclinical hypothyroidism on pregnancy outcomes: a prospective multicenter single-blind, randomized, controlled study of thyroid function screening test during pregnancy. *J Matern Fetal Neonatal Med* 2016;29:1391–4.
22. Nazarpour S, Ramezani Tehrani F, Simbar M, *et al*. Effects of levothyroxine treatment on pregnancy outcomes in pregnant women with autoimmune thyroid disease. *Eur J Endocrinol* 2017;176:253–65.
23. Negro R, Schwartz A, Gismondi R, *et al*. Universal screening versus case finding for detection and treatment of thyroid hormonal dysfunction during pregnancy. *J Clin Endocrinol Metab* 2010;95:1699–707.
24. Taylor P L. Controlled antenatal thyroid screening (CATS) study: Obstetric outcomes. *European Thyroid Journal*;5:132. Conference:39th Annual Meeting of the European Thyroid Association, ETA. 2016. Denmark. Conference Start: 20160903. Conference End: 06.
25. Varner M. 70: Thyroid function in neonates of women with subclinical hypothyroidism and hypothyroxinemia. *Am J Obstet Gynecol* 2017;216:S49–S50. Conference: 37th annual meeting of the society for maternal-fetal medicine: the pregnancy meeting United states. Conference start: 20170123;Conference end: 20170128 216.
26. Hales C, Taylor PN, Channon S, *et al*. Controlled antenatal thyroid screening II: effect of treating maternal suboptimal thyroid function on child cognition. *J Clin Endocrinol Metab* 2018;103:1583–91.
27. Liu H, Shan Z, Li C, *et al*. Maternal subclinical hypothyroidism, thyroid autoimmunity, and the risk of miscarriage: a prospective cohort study. *Thyroid* 2014;24:1642–9.
28. Stagnaro-Green A. Second trimester levothyroxine treatment for subclinical hypothyroidism or hypothyroxinaemia of pregnancy does not improve cognitive outcomes of children. *Evid Based Med* 2017;22:149.
29. Vissenberg R, van Dijk MM, Fliers E, *et al*. Effect of levothyroxine on live birth rate in euthyroid women with recurrent miscarriage and TPO antibodies (T4-LIFE study). *Contemp Clin Trials* 2015;44:134–8.
30. Coomarasamy A. Thyroid AntiBodies and LEvoThyroxine study (TABLET). 2011. http://www.isrctn.com/ISRCTN15948785
31. Wang H, Gao H, Chi H, *et al*. Effect of levothyroxine on miscarriage among women with normal thyroid function and thyroid autoimmunity undergoing in vitro fertilization and embryo transfer: a randomized clinical trial. *JAMA* 2017;318:2190.
32. Korevaar TIM, Pop VJ, Chaker L, *et al*. Dose Dependency and a Functional Cutoff for TPO-Antibody Positivity During Pregnancy. *J Clin Endocrinol Metab* 2018;103:778–89.
33. Casey B. 65: Thyroid peroxidase antibodies in women with subclinical hypothyroidism or hypothyroxinemia. *Am J Obstet Gynecol* 2017;216(1):S47. Conference start: 20170123;Conference end: 20170128 216.
34. Plowden TC, Schisterman EF, Sjaarda LA, *et al*. Subclinical Hypothyroidism and Thyroid Autoimmunity Are Not Associated With Fecundity, Pregnancy Loss, or Live Birth. *J Clin Endocrinol Metab* 2016;101:2358–65.
35. Thompson W, Russell G, Baragwanath G, *et al*. Maternal thyroid hormone insufficiency during pregnancy and risk of neurodevelopmental disorders in offspring: A systematic review and meta-analysis. *Clin Endocrinol* 2018;88:575–84.
36. Li J, Shen J, Qin L. Effects of Levothyroxine on Pregnancy Outcomes in Women With Thyroid Dysfunction: A Meta-analysis of Randomized Controlled Trials. *Altern Ther Health Med* 2017;23:49–58.