

Levothyroxine treatment and pregnancy outcome in women with subclinical hypothyroidism undergoing assisted reproduction technologies: systematic review and meta-analysis of RCTs

B. Velkeniers¹, A. Van Meerhaeghe², K. Poppe¹, D. Unuane¹,
H. Tournaye³, and P. Haentjens^{4,5,*}

¹Department of Endocrinology and General Internal Medicine, Universitair Ziekenhuis Brussel, Vrije Universiteit Brussel, Brussels, Belgium
²GERHPAC (Group of Applied Epistemology and Rational Clinic of the Public Hospitals from Charleroi) ISPPC - CHU-Charleroi, Charleroi, Belgium
³Centre for Reproductive Medicine, Universitair Ziekenhuis Brussel, Vrije Universiteit Brussel, Brussels, Belgium
⁴Center for Outcomes Research and Laboratory for Experimental Surgery, Universitair Ziekenhuis Brussel, Vrije Universiteit Brussel, Laarbeeklaan 101, B-1090 Brussels, Belgium
⁵CEBAM Belgian Center for Evidence Based Medicine, Belgian Branch of the Cochrane Collaboration, Leuven, Belgium

*Correspondence address. Tel: +32-2-477-64-18; Fax: +32-2-477-64-28; E-mail: patrick.haentjens@uzbrussel.be

Submitted on September 26, 2012; resubmitted on November 2, 2012; accepted on November 14, 2012

TABLE OF CONTENTS

- Introduction
- Methods
- Results
 - Trial characteristics
 - Delivery (primary analyses)
 - Other outcomes (secondary analyses)
- Discussion
- Conclusions

BACKGROUND: Previous meta-analyses of observational data indicate that pregnant women with subclinical hypothyroidism have an increased risk of adverse pregnancy outcome. Potential benefits of levothyroxine (LT4) supplementation remain unclear, and no systematic review or meta-analysis of trial findings is available in a setting of assisted reproduction technologies (ART).

METHODS: Relevant trials published until August 2012 were identified by searching MEDLINE, EMBASE, Web of Knowledge, the Cochrane Controlled Trials Register databases and bibliographies of retrieved publications without language restrictions.

RESULTS: From 630 articles retrieved, we included three trials with data on 220 patients. One of these three trials stated 'live delivery' as outcome. LT4 treatment resulted in a significantly higher delivery rate, with a pooled relative risk (RR) of 2.76 (95% confidence limits 1.20–6.44; $P = 0.018$; $I^2 = 70\%$), a pooled absolute risk difference (ARD) of 36.3% (3.5–69.0%; $P = 0.030$) and a summary number needed to treat (NNT) of 3 (1–28) in favour of LT4 supplementation. LT4 treatment significantly lowered miscarriage rate with a pooled RR of 0.45 (0.24–0.82; $P = 0.010$; $I^2 = 26\%$), a pooled ARD of –31.3% (–48.2 to –14.5%; $P < 0.001$) and a summary NNT of 3 (2–7) in favour of LT4 supplementation. LT4 treatment had no effect on clinical pregnancy (RR 1.75; 0.90–3.38; $P = 0.098$; $I^2 = 82\%$). In an ART setting, no data are available on the effects of LT4 supplementation on premature delivery, arterial hypertension, placental abruption or pre-eclampsia.

CONCLUSIONS: Our meta-analyses provide evidence that LT4 supplementation should be recommended to improve clinical pregnancy outcome in women with subclinical hypothyroidism and/or thyroid autoimmunity undergoing ART. Further research is needed to determine pregnancy outcome after close monitoring of thyroid function to maintain thyroid-stimulating hormone and free T4 levels within the trimester-specific reference ranges for pregnancy.

Key words: subclinical hypothyroidism / thyroid autoimmunity / assisted reproduction technologies / pregnancy outcome / levothyroxine

Introduction

Previous meta-analyses of observational data indicate that pregnant women with subclinical hypothyroidism have an increased risk of adverse pregnancy outcome (Toulis *et al.*, 2010; van den Boogaard *et al.*, 2011). Potential benefits of levothyroxine (LT4) supplementation remain unclear (Reid *et al.*, 2010; Thangaratinam *et al.*, 2012; Vissenberg *et al.*, 2012), mainly because the evidence from randomized controlled trials (RCTs) within a setting of assisted reproduction technologies (ART) is sparse and conflicting. Hence, no systematic review or meta-analysis of trial findings is available in an ART setting.

In women of reproductive age, thyroid dysfunction and thyroid autoimmunity (TAI) are prevalent. Also, thyroid dysfunction is frequently associated with TAI (van den Boogaard *et al.*, 2011). Our group has documented that in infertile women the prevalence of those positive for thyroid peroxidase antibodies (TPO-Abs) is higher than in fertile controls, especially in the presence of polycystic ovarian disease and endometriosis (Poppe *et al.*, 2002). Both thyroid dysfunction and TAI have independently been associated with adverse fertility and pregnancy outcomes. In two recently published meta-analyses, the presence of TAI was associated with an increased risk for spontaneous miscarriage in subfertile women achieving pregnancy through an IVF procedure (Toulis *et al.*, 2010; Thangaratinam *et al.*, 2012). (Sub)clinical hypothyroidism increases the odds of pregnancy complications, including pre-eclampsia, placental abruption, preterm birth and neonatal mortality (van den Boogaard *et al.*, 2011). Evidence is far clearer for overt than for subclinical hypothyroidism and, although a systematic review has noted these effects in subclinical hypothyroidism, effects observed in individual studies have been quite variable (Casey *et al.*, 2007; Cleary-Goldman *et al.*, 2008). All meta-analyses focused on adverse events rather than benefits of intervention such as delivery, and did not report summary data specifically pertaining to an ART setting, for example, implantation rate or clinical pregnancy rate. Adverse outcomes and beneficial impact of treatment of thyroid disorders on fertility and subsequent pregnancy may critically determine the necessity to screen and treat women of infertile couples (Vissenberg *et al.*, 2012).

Our purpose is to summarize the evidence from RCTs which address the following research question: 'In subfertile women with subclinical hypothyroidism and TAI undergoing ART, what is the effect of LT4 supplementation, compared with placebo or no treatment, on pregnancy outcome—including, clinical pregnancy, miscarriage, delivery, premature delivery, arterial hypertension, pre-eclampsia and placental abruption?'

Methods

We conducted a systematic review of English and non-English articles using MEDLINE (Ovid, PubMed), EMBASE, Web of Knowledge and the

Cochrane Controlled Trials Register databases, with the last computerized searches undertaken in August 2012. The search terms, based on patient population, setting, intervention and outcomes, were as follows: 'thyroid, hypothyroidism, subclinical hypothyroidism, thyroid autoimmunity, thyroperoxidase antibodies (TPO-Abs), thyroglobulin antibodies (Tg-Abs)' AND 'assisted reproductive technologies, *in vitro* fertilisation (IVF), intracytoplasmic sperm injection (ICSI), ovarian stimulation AND levothyroxine' AND 'pregnancy outcome, delivery, miscarriage, preterm birth'. The computerized searches were supplemented by a manual search of the bibliographies of all retrieved articles and the authors' files. Potentially relevant articles were assessed for inclusion against prespecified eligibility and exclusion criteria.

We included RCTs, published in full that assessed the effect of LT4 supplementation, compared with placebo or no treatment, on pregnancy outcome in subfertile women with subclinical hypothyroidism and/or TAI undergoing ART. Our primary outcome of interest was delivery. Therefore, studies were required to report delivery as a separate outcome and to detail the number of confirmed events in all treatment arms. Reviews, abstracts, research letters, observational cohort or case-control studies, uncontrolled studies and the studies in which the number of delivery events was not reported in all treatment arms were excluded. We also excluded RCTs not conducted in a setting of ART.

Three investigators (P.H., B.V. and A.V.M.) independently collected data on patient and study characteristics, treatment interventions and clinical outcomes based on published information only without contacting researchers to collect additional data. Assessment of the methodological quality of the trials was based on the Jadad (Jadad *et al.*, 1996) and PEDro scoring systems (available at <<http://www.pedro.org.au/english/downloads/pedro-scale/>> accessed 12 November 2012). Our primary outcome event was delivery (live birth or delivery as reported in the original papers). Secondary outcome events (yes/no; binary outcome data) included implanting embryos, clinical pregnancy, miscarriage, premature delivery, arterial hypertension, pre-eclampsia and placental abruption. We also extracted information on the number of oocytes retrieved, mature oocytes, fertilized oocytes, embryos transferred and embryos cryopreserved (continuous outcome data). Data were independently extracted by two authors (P.H. and B.V.) and checked for accuracy in a second review. Consensus was achieved for all data.

For each individual trial and for each outcome of interest, we computed an effect size and its 95% confidence interval (CI). The effect sizes were the relative risk (RR), the absolute risk difference (ARD) and the number needed to treat (NNT) for binary outcome data, and the standardized mean difference (SMD) for continuous outcome data (Borenstein *et al.*, 2009).

For each outcome of interest, the effect sizes of the individual trials were pooled using DerSimonian and Laird random-effects models (Borenstein *et al.*, 2009). Random-effect models assume that the observed variability between the studies and their studied populations reflects sampling variability and heterogeneity of the study populations (Borenstein *et al.*, 2009).

The results were examined for heterogeneity by visually examining forest plots, by using the Cochran Q test, and by computing the I^2 statistic,

with values <25% indicating low, 25–50% indicating moderate and values exceeding 50% indicating high heterogeneity.

To evaluate the effect of each selected study on the overall results of the meta-analysis we decided, *a priori*, to perform a one-way sensitivity analysis for our primary outcome. To explain anticipated heterogeneity among trial findings, we identified, also *a priori*, potential sources of heterogeneity, including cut-off levels used for the diagnosis of thyroid disorder [especially cut-offs for thyroid-stimulating hormone (TSH) values], women's ethnicity, women's mean age, causes of infertility, ART procedure, total sample size, year of trial publication and withdrawal rates.

Publication bias was explored visually by the funnel plot method, and formally quantified by the classic fail-safe N and Duval and Tweedie's trim-and-fill methods (Borenstein *et al.*, 2009). The classic fail-safe N method computes the number of missing trials that would bring the *P*-value to larger than alpha. Duval and Tweedie's trim-and-fill method looks for missing studies on the funnel plot and recalculates an adjusted pooled effect size by also including the potentially missed (trimmed) studies. There must be at least three studies published to run this publication bias procedure.

Statistical processes for combining data from multiple trials were performed in CMA, version 2 (Comprehensive Meta-Analysis, Biostat TM, Englewood, NJ, USA).

Results

Trial characteristics

Figure 1 shows the selection process after the search. Only three RCTs met our inclusion criteria (Fig. 1, Table I). The three full papers addressing our research question were published from 2005 to 2011 and provided data on 220 women undergoing ART in Italy (Negro *et al.*, 2005), Egypt (Abdel Rahman *et al.*, 2010) and South Korea (Kim *et al.*, 2011). Thyroid disorders were diagnosed based on the presence of TPO-Abs in one trial (Negro *et al.*, 2005), and on an increased TSH value, using a cut-off level of 4.0 or 4.5 mIU/l, in both other trials (Abdel Rahman *et al.*, 2010; Kim *et al.*, 2011). Causes of infertility and ART procedures were comparable among trials, although the Negro *et al.* (2005) trial did not include male factor infertility and the Kim *et al.* (2011) trial did not include patients with ovarian dysfunction. LT4 supplementations were dissimilar, some being individually tailored to TSH values, others using fixed doses of LT4. Recommended TSH cut-off values before pregnancy (<2.5 mIU/l according to Endocrine Society clinical practice guideline, De Groot *et al.*, 2012) were not always available in the trial

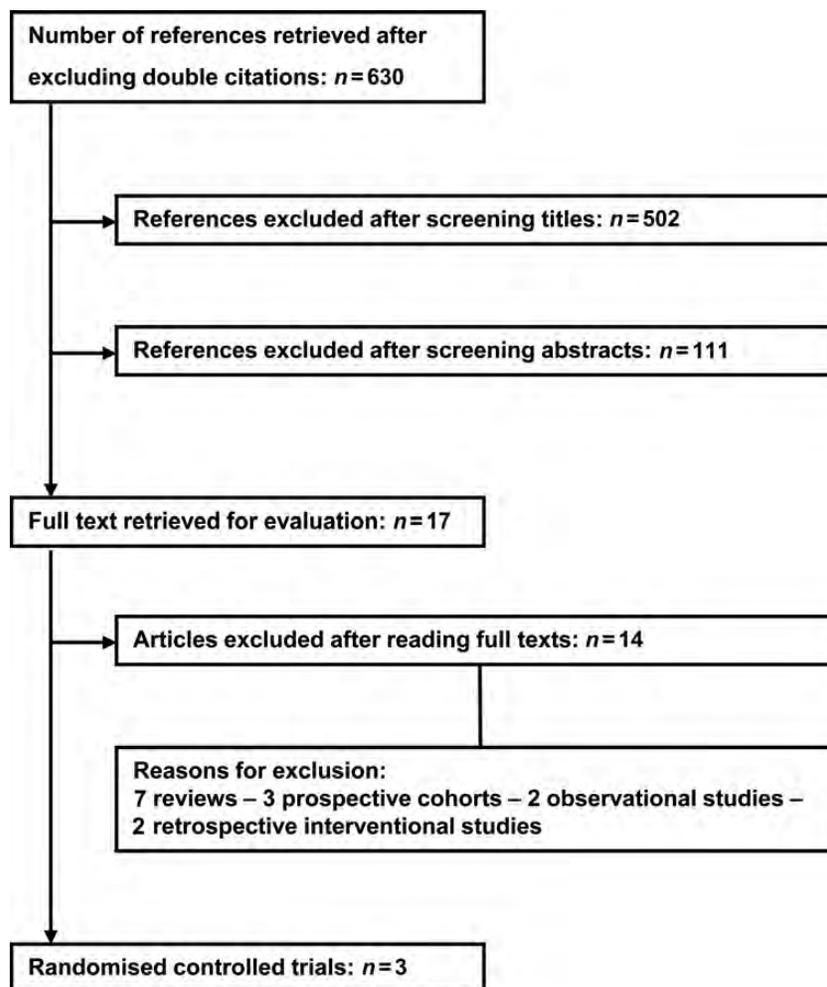


Figure 1 PRISMA flow chart.

Table I-a Characteristics of the 3 randomized clinical trials included in the primary analyses

RCT	Country	Population	Causes of infertility, according to treated and control groups of trial	Intervention
Negro et al. (2005)	Italy	86 TPO-Abs-positive infertile women undergoing IVF/ICSI	Ovarian dysfunction, <i>n</i> (%), treated 11 (31), placebo 13 (36) Tubal factors, <i>n</i> (%), treated 10 (28), placebo 9 (25) Endometriosis, <i>n</i> (%), treated 7 (19), placebo 9 (25) Idiopathic, <i>n</i> (%), treated 8 (22), placebo 5 (14)	Levothyroxine 1 µg/day
Abdel Rahman et al. (2010)	Egypt	70 infertile women with subclinical hypothyroidism undergoing IVF/ICSI	Ovarian dysfunction, <i>n</i> (%), treated 13 (37), placebo 12 (34) Tubal factors, <i>n</i> (%), treated 9 (26), placebo 11 (31) Endometriosis, <i>n</i> (%), treated 7 (20), placebo 5 (14) Idiopathic, <i>n</i> (%), treated 6 (17), placebo 7 (20)	Levothyroxine 50 to 100 µg/day to normalize TSH before ART
Kim et al. (2011)	South Korea	64 infertile women with subclinical hypothyroidism undergoing IVF/ICSI	Tubal factor, <i>n</i> (%), treated 10 (31), not treated 9 (28) Endometriosis III or IV, <i>n</i> (%), treated 6 (19), not treated 7 (22) Male factor, <i>n</i> (%), treated 13 (41), not treated 13 (41) Unexplained, <i>n</i> (%), treated 3 (9), not treated 3 (9)	Levothyroxine 50 µg/day before pregnancy; Titration during pregnancy to maintain TSH < 2.5 mIU/L

Table I-b Trial characteristics (cont'd)

RCT	Reference values for thyroid status	Definition of thyroid status and thyroid hormone values	Haddad quality score	PEDro quality score
Negro et al. (2005)	TSH 0.27–4.2 mIU/L fT4 9.3–18.0 ng/L or 12–33.5 pmol/L TPO-Ab 0–100 kIU/L	values within normal limits and controlled before ART procedure, after treatment mean (SD) for TSH: 1.6 (0.8) mean (SD) for fT4: 12.0 (2.0) mean (SD) baseline TSH: 1.9 (0.7) before treatment, 1.1 (0.3) after treatment; not treated 1.7 (0.7) mean (SD) baseline fT4: 11.2 (1.8) before treatment, 14.1 (2.5) after treatment; not treated 11.7 (2.1)	5 / 5 excellent	10 / 11
Abdel Rahman et al. (2010)	TSH 0.27–4.2 mIU/L fT4 9.0–25.9 ng/L or 0.9-2.59 ng/dL	TSH > 4mUI/L, fT4 within normal range and controlled before start ART mean (SD) baseline TSH before treatment: placebo 4.8 (0.7), treated 4.7 (0.5) mean (SD) baseline fT4 before treatment: placebo 1.04 (0.49), treated 1.0 (0.4) mean (SD) after treatment: TSH placebo 4.9 (0.3), treated 1.1 (0.7); fT4 placebo 1.01 (0.5), treated 0.95 (0.4)	5 / 5 excellent	10 / 11
Kim et al. (2011)	TSH 0.27–4.0 mIU/L fT4 0.9–2.59 ng/dL	TSH > 4.5 mUI/L, fT4 within normal range and controlled on the day of beta-HCG measurement mean (SD) baseline TSH before treatment : placebo 6.7 (1.8), treated 6.6 (1.7) mean (SD) baseline fT4 before treatment: placebo 1.2 (0.2); treated 1.2 (0.2) mean (SD) after treatment on day of beta-HCG measurement: TSH placebo 6.9 (2.0), treated 2.3 (0.4); fT4 placebo 1.0 (0.2), treated 1.4 (0.3)	3 / 5 good	8 / 11

publication. Two trials (Negro *et al.*, 2005; Abdel Rahman *et al.*, 2010) used the third-generation electrochemiluminescence immunoassay for TSH (Roche, Germany); one trial (Kim *et al.*, 2011) used an immunoradiometric TSH assay (Diasorin, Saluggia, Italy). Two trials were placebo-controlled (Negro *et al.*, 2005; Abdel Rahman *et al.*, 2010). One had a pragmatic open design (Kim *et al.*, 2011). Only one paper clearly reported live birth delivery (Kim *et al.*, 2011); the two other papers reported delivery without further specifying this event (Negro *et al.*, 2005; Abdel Rahman *et al.*, 2010). All analyses were conducted on an intention-to-treat basis. None of the papers presented results of between-group statistical comparisons using a well-known effect size, such as an odds ratio, an RR, an absolute between group difference or a NNT, even if a statistically significant difference was observed (Table I; PEDro score). Drop-out rates were zero in all trials.

Delivery (primary analyses)

LT4 treatment resulted in a significantly higher delivery rate, with a pooled RR of 2.76 (95% confidence limits 1.20–6.44; $P = 0.018$; Cochrane's P -value for heterogeneity 0.034, $I^2 = 70\%$), a pooled ARD of 36.3% (3.5–69.0%; $P = 0.030$) and a summary NNT to gain one additional delivery of 3 (1–28) in favour of LT4 supplementation (Fig. 2, forest plot; Table II, primary end-point). The high and statistically significant heterogeneity could be explained by non-overlapping CIs of individual studies (Fig. 2, forest plot).

One-way sensitivity analyses demonstrated that the overall effect size and its statistical significance were consistent across the studies and did not depend on any single study (data not shown). Interestingly, I^2 values ranged from 0% (Abdel Rahman *et al.*, trial omitted) to 90% (Kim *et al.*, trial omitted), suggesting that high I^2 values were related to no overlapping CIs of individual trials. Indeed, in the univariate sensitivity analysis, none of the *a priori* variables were significant: type of

control (placebo versus no treatment; $P = 0.75$), total sample size ($P = 0.63$), participants' mean age ($P = 0.97$) and year of trial publication ($P = 0.51$). We were unable to perform the other sensitivity analyses, also defined *a priori*, given identical (e.g. causes of infertility, type of ART procedure, drop-out rates) or too dissimilar trial characteristics not allowing to define specific categories (e.g. cut-off levels used for the diagnosis of thyroid disorder, women's ethnicity, LT4 treatment schedule).

The funnel plot provided no evidence of publication bias. The fail-safe N number (12 missing studies needed to bring P -value larger than 0.05) and the trim-and-fill results (no trimmed studies; identical pooled effect sizes and 95% CIs) indicated that publication bias was unlikely to affect our findings.

Other outcomes (secondary analyses)

All three trials reporting delivery data also provided data on other outcomes of interest. The results of these secondary analyses are summarized in Table II, secondary end-point. LT4 treatment significantly lowered miscarriage rate with a pooled RR of 0.45 (0.24–0.82; $P = 0.010$; Cochrane's P -value for heterogeneity 0.26; $I^2 = 26\%$), a pooled ARD of -31.3% (-48.2 to -14.5% ; $P < 0.001$) and a summary NNT to prevent one miscarriage of 3 (2–7) in favour of LT4 supplementation. Heterogeneity was low. The funnel plot was asymmetric. The fail-safe N number, with six missing studies needed to bring P -value larger than 0.05, indicated that publication bias was unlikely to affect our findings. The trim-and-fill results, on the other hand, with one trimmed study to right of the mean but with similar 95% CIs still not including the value of no effect, suggested that publication bias was unlikely to materially affect these findings.

LT4 treatment resulted in a significantly higher number of fertilized oocytes (two trials; SMD 0.55, 0.03–1.08; $P = 0.039$) and implantation rate (one trial; RR 1.81, 1.01–3.25; $P = 0.049$; ARD 12.0%,

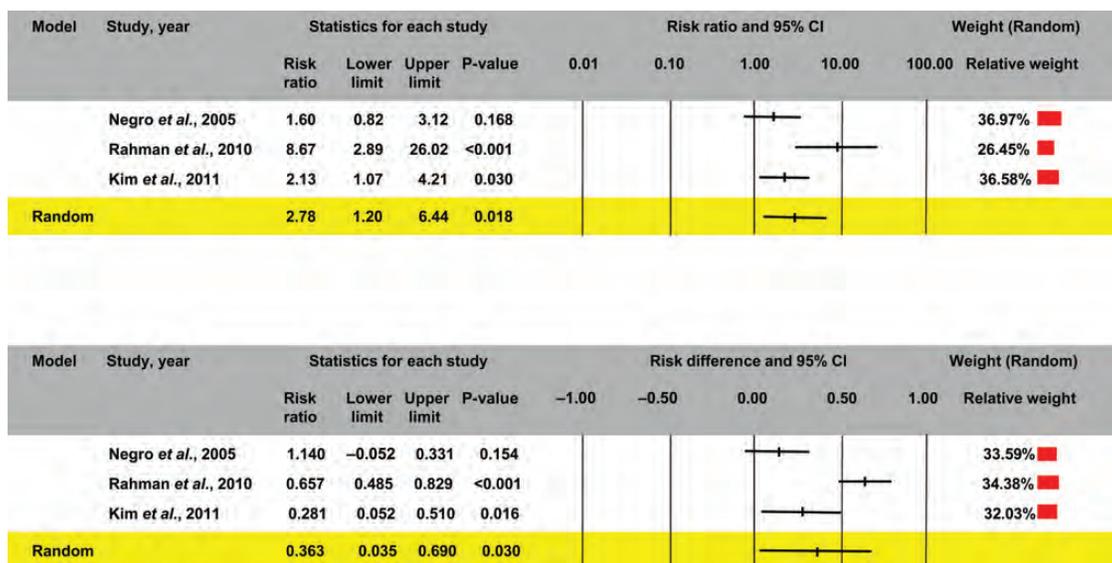


Figure 2 Forest plots comparing delivery outcome events between LT4 supplementation and placebo/no treatment groups for the following random-model effect sizes: RR (upper panel) and ARD (Risk difference, lower panel). Point estimates to the right of the vertical line of no effect (1.00 for RR, 0.00 for risk difference) indicate an effect in favour of LT4 treatment. CI, confidence interval.

Table II Summary of outcomes and effect sizes for the primary and secondary analyses.

Outcome	Number of studies	Number of participants	Statistical method	Effect size, calculated	I^2
Primary end-point (primary analyses)					
Delivery	3	220	RR (Random, 95% CI)	2.76 [1.20, 6.44]	70%
Delivery	3	220	ARD (Random, 95% CI)	36.3% [3.5%, 69.0%]	88%
Delivery	3	220	NNT (Random, 95% CI)	3 [1, 28]	NA
Secondary end-points (secondary analyses)					
Oocytes retrieved	2	134	SMD (Random, 95% CI)	0.08 [−0.26, 0.42]	0%
Mature oocytes	2	134	SMD (Random, 95% CI)	0.64 [−0.20, 1.46]	78%
Fertilized oocytes	2	134	SMD (Random, 95% CI)	0.55 [0.03, 1.08]	47%
Embryos transferred	1	64	SMD (Random, 95% CI)	0.00 [−0.22, 0.22]	NA
Embryos cryopreserved	1	64	SMD (Random, 95% CI)	−0.70 [−0.53, 1.93]	NA
Embryo implantation	1	64	RR (Random, 95% CI)	1.81 [1.01, 3.25]	NA
Embryo implantation	1	64	ARD (Random, 95% CI)	12.0% [0.5%, 23.5%]	NA
Embryo implantation	1	64	NNT (Random, 95% CI)	9 [5, 200]	NA
Clinical pregnancy	3	220	RR (Random, 95% CI)	1.75 [0.90, 3.38]	82%
Clinical pregnancy	3	220	ARD (Random, 95% CI)	32.0% [−12.1, 76.0]	93%
Clinical pregnancy	3	220	NNT (Random, 95% CI)	NA	NA
Miscarriage	3	119	RR (Random, 95% CI)	0.45 [0.24, 0.82]	26%
Miscarriage	3	119	ARD (Random, 95% CI)	−31.3% [−48.2%, −14.5%]	0%
Miscarriage	3	119	NNT (Random, 95% CI)	3 [2, 7]	NA

NA, not applicable; NA for effect size column means, NNT not calculated as difference not statistically significant; NA for I^2 column means, one trial only.

For RRs values of >1 are in favour of LT4 treatment; for SMDs and ARDs, values >0 are in favour of LT4 treatment.

95% CIs excluding the value of no effect (1 for RRs; 0 for differences) are printed bold.

0.5–23.5%; $P = 0.042$); NNT to gain one additional implanted embryo 9 (5–200).

LT4 treatment had no effect on clinical pregnancy rate (three trials pooled RR $P = 0.098$; pooled ARD $P = 0.156$), number of retrieved oocytes (two trials; pooled SMD $P = 0.651$), number of mature oocytes (two trials; pooled SMD $P = 0.136$), number of embryos transferred (one trial; between-group difference $P = 1.00$) and number of embryos cryopreserved (one trial; $P = 0.264$; Table II).

In an ART setting, no data were available on premature delivery, arterial hypertension, placental abruption or pre-eclampsia.

Discussion

In the meta-analysis reported here, LT4 supplementation versus placebo/no treatment demonstrated a significant increase in delivery, the primary outcome, and implantation of embryos, together with a significant decrease in miscarriage. The pooled risk difference indicates that in order to gain one delivery, only three women with subclinical hypothyroidism and/or TAI undergoing ART need to be supplemented with LT4. To prevent one miscarriage, only three women need to be treated. The corresponding NNT for embryo implantation was 9.

At the time of writing, five systematic reviews and meta-analyses reported on the association between thyroid dysfunction and pregnancy outcome and the impact of interventions (Reid et al., 2010; Toulis et al., 2010; van den Boogaard et al., 2011; Thangaratinam et al., 2012; Vissenberg et al., 2012). These meta-analyses focused on adverse events rather than benefits of intervention, for example,

a potential increase in delivery events. These meta-analyses did not report summary data specifically pertaining to an ART setting, such as clinical pregnancy, oocytes retrieved, mature oocytes, fertilized oocytes, embryos transferred, embryos cryopreserved and embryo implantation. Three of these five meta-analyses also included data from observational studies, including prospective cohorts, retrospective cohorts and case-control designs. Our analyses, on the other hand, relied on trial evidence within the context of ART. This approach enables us to present summary findings in terms of RRs, ARDs and NNT, and also to extend the findings of previously published meta-analyses by providing quantified information on end-points pertaining to this setting.

Our review and meta-analyses cannot rule out potential bias secondary to non-reporting of negative study results. Several outcome events were reported in one single trial only. We could not analyse time-related improvements in medical care or potential influences of ethnicity and other confounding factors given the small number of trials and their identical or very dissimilar trial characteristics. Thyroid function during subsequent pregnancy may have been different among studies but follow-up values of thyroid function tests during pregnancy were not always reported in the primary trials.

Our analyses indicated that the point estimates of effect sizes of individual trials reporting on the outcome of pregnancy rate after ART were consistent for the direction of the effect, i.e. all suggested a beneficial effect in favour of LT4 treatment. Differences in magnitude of effect between trials and high heterogeneity observed in the current meta-analyses may relate to the immunoassays used and the cut-off

values used to define the degree of thyroid dysfunction, the presence of thyroid antibodies in the selected patient population undergoing an ART procedure, differences in LT4 supplementation schedules in the treatment arms and contamination by LT4 supplementation in the control arms.

In the Negro trial (Negro *et al.*, 2005), which included TPO-Abs-positive euthyroid women, LT4 did not significantly alter delivery rate compared with placebo. Thyroid tests were checked only once, more precisely 1 month after starting LT4 treatment but before the ART procedure. A fixed dose of LT4 was used therefore one cannot assume that target values of <2.5 mIU/l were met during ART and pregnancy in all patients. In the Abdel Rahman trial (Abdel Rahman *et al.*, 2010), which included patients with subclinical hypothyroidism, the dose of LT4 was individually adjusted to obtain normal values of TSH before starting the ART procedure. Given these inclusion criteria, patients had a worse baseline thyroid function and therefore one could expect a larger effect of treatment with LT4. If and to what extent the presence of TAI might influence the effect size cannot be inferred because TPO-Abs were not measured in this trial. Finally, the Kim trial (Kim *et al.*, 2011) included women with subclinical hypothyroidism based on thyroid function tests. TAI was also evaluated. In the control group both TPO-Abs (thyroid peroxidase antibodies) and Tg-Abs (thyroglobulin antibodies) levels were significantly higher in the subgroup with miscarriage versus the subgroup with delivery; levels were similar in the treated group. In the treatment arm LT4 treatment was initially 50 µg, but during subsequent pregnancy the doses were adjusted to reach target TSH values of <2.5 mIU/l in first trimester of pregnancy. Patients developing overt hypothyroidism during pregnancy in the control arm were also prescribed LT4. The difference in effect size obtained in this trial compared with the Rahman trial may be partly explained both by imbalances in TPO-Abs and Tg-Abs levels and by contamination resulting from LT4 supplementation in the control arm.

Our results supporting LT4 supplementation for the prevention of miscarriage/improvement of both embryo implantation and delivery after ART are underpinned by a strong biological rationale.

Both low thyroid hormone levels (hypothyroidism) and/or the presence of TPO-Abs (thyroid autoimmunity) are associated with infertility (Poppe *et al.*, 2002; Negro *et al.*, 2006; Unuane *et al.*, 2011).

Clinical (overt) hypothyroidism and subclinical hypothyroidism have been linked with poor reproductive outcome (Unuane *et al.*, 2011; van den Boogaard *et al.*, 2011). In one observational study, TSH levels were inversely proportional to the fertilization rate after ART (Cramer *et al.*, 2003). These findings, however, were not reproduced in a recent retrospective study (Michalakis *et al.*, 2011), where TSH values within the range of 2.5–4 mIU/l were considered moderately elevated. The prevalence of moderate high baseline TSH values measured within 6 weeks before ovarian stimulation was significantly higher (23%) compared with the prevalence reported in previous studies (Michalakis *et al.*, 2011) because the new cut-off values in accordance with the guidelines of the Endocrine Society were used. Neither the presence/absence of thyroid antibodies nor the changes in thyroid hormone levels during the ART procedure were reported, even though elevated TSH values of >4 mIU/l before ART were significantly associated with a diminished ovarian reserve.

Overt hypothyroidism has long since been associated with menstrual disorders. Treatment of hypothyroidism with LT4 usually

restores menstrual pattern and reverses hormonal alterations, and may thus improve fertility (Krassas *et al.*, 2010). We identified no data on menstrual abnormalities in subclinical hypothyroidism. Our finding of better fertilization rates among women treated with LT4 are in line with previous observational data from Scoccia *et al.* (2012) reporting lower oocyte fertilization rate in hypothyroid women compared with euthyroid women. The reason for these findings remains largely speculative. TSH and thyroid hormone receptors have been located in human granulosa cells and both tri-iodothyronine (T3) and thyroxine (T4) are found in the follicular fluid. Alterations in TSH levels may therefore negatively influence oocyte quality and function in (sub)clinical hypothyroid patients.

There are two major hypotheses on the causal pathway between the presence of TPO-Abs (thyroid autoimmunity) and fertility and obstetric complications. The first hypothesis is that autoimmunity increases the risk for developing (sub)clinical hypothyroidism, especially in the setting of an ART procedure (Poppe *et al.*, 2008). The second hypothesis is that thyroid antibodies can be considered an expression of autoimmunity in general and that the adverse fertility and obstetric outcomes may be caused by a related immune dysfunction and/or associated general autoimmune disorders, for example, the presence of anticardiolipin antibodies (Toulis *et al.*, 2010).

The debate on the contribution of TAI versus thyroid dysfunction in patients with subclinical thyroid dysfunction remains unsettled. Yet, intervention trials with LT4 in patients with TAI and/or subclinical hypothyroidism may drive the balance towards thyroid dysfunction as a major contributor to worse fertility and pregnancy outcomes. It is therefore important to quantify the effects of the LT4 intervention in the setting of an ART procedure.

Overall, the current meta-analyses have been performed to gain insights into the effect that may be expected by an intervention using LT4 in patients with TAI and/or subclinical hypothyroidism. From our data, we can conclude that patients with both the presence of TAI and/or of subclinical thyroid dysfunction may benefit from this treatment, with a low NNT. The primary outcome measure was delivery (pregnancy outcome) but surrogate end-points, including number and quality of oocytes, embryo implantation rate and hence pregnancy rate after ART, were also included in our analyses as secondary outcome measures. These end-points might provide better insights pertaining to the quality of life of patients, in terms of the number of ART procedures—end-points not captured by the trials and the current meta-analyses, even though not negligible in this patient population.

Finally, the benefit of treatment should be weighed against possible side effects. Treatment should be monitored closely to avoid over-treatment, as documented in patients receiving LT4 outside of pregnancy (Canaris *et al.*, 2000).

Randomized placebo-controlled trials with LT4 in infertile women with TAI and/or subclinical hypothyroidism are warranted to further study pregnancy outcomes after ART, using close monitoring of thyroid function in order to maintain TSH and free T4 levels within the trimester-specific reference ranges for pregnancy.

Conclusions

In conclusion, our meta-analyses of RCTs confirm and extend the findings from previous primary research and meta-analyses, indicating that

patients with TAI and subclinical hypothyroidism who are undergoing ART may benefit from a simple intervention with LT4 supplementation in order to improve fertility and subsequent pregnancy outcomes (delivery).

Author's roles

B.V. had contributed to the idea for the study. P.H. and B.V. designed the study. P.H., B.V. and A.V.M. collected the data, had full access to all of the data, take responsibility for the integrity of the data and the accuracy of the data analysis, and are the guarantors for the paper. P.H. did the statistical analysis. H.T., K.P., D.U. and A.V.M. provided advice on interpretation of the results. P.H., B.V. and A.V.M. drafted the paper. All authors revised the paper critically and approved the final manuscript.

Funding

This work was supported by a grant from the Willy Gepts Foundation, Universitair Ziekenhuis Brussel, to Dr Brigitte Velkeniers.

Conflict of interest

None declared.

References

- Abdel Rahman AH, Aly Abbassy H, Abbassy AA. Improved in vitro fertilization outcomes after treatment of subclinical hypothyroidism in infertile women. *Endocr Pract* 2010;**16**:792–797. Erratum in *Endocr Pract* 2011;**17**:526.
- Borenstein M, Hedges L, Higgins J, Rothstein H. *Introduction to Meta-Analysis*. Hoboken, NJ: Wiley, 2009.
- Canaris GJ, Manowitz NR, Mayor G, Ridgway EC. The Colorado thyroid disease prevalence study. *Arch Intern Med* 2000;**160**:526–534.
- Casey BM, Dashe JS, Spong CY, McIntire DD, Leveno KJ, Cunningham GF. Perinatal significance of isolated maternal hypothyroxinemia identified in the first half of pregnancy. *Obstet Gynecol* 2007;**109**:1129–1135.
- Cleary-Goldman J, Malone FD, Lambert-Messerlian G, Sullivan L, Canick J, Porter TF, Luthy D, Gross S, Bianchi DW, D'Alton ME. Maternal thyroid hypofunction and pregnancy outcome. *Obstet Gynecol* 2008;**112**:85–92.
- Cramer DW, Sluss PM, Powers RD, McShane P, Ginsburgs ES, Hornstein MD, Vitonis AF, Barbieri RL. Serum prolactin and TSH in an in vitro fertilization population: is there a link between fertilization and thyroid function? *J Assist Reprod Genet* 2003;**20**:210–215.
- De Groot L, Abalovich M, Alexander EK, Amino N, Barbour L, Cobin RH, Eastman CJ, Lazarus JH, Luton D, Mandel SJ et al. Management of thyroid dysfunction during pregnancy and postpartum: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab* 2012;**97**:2543–2565.
- Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, McQuay HJ. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials* 1996;**17**:1–12.
- Kim CH, Ahn JW, Kang SP, Kim SH, Chae HD, Kang BM. Effect of levothyroxine treatment on in vitro fertilization and pregnancy outcome in infertile women with subclinical hypothyroidism undergoing in vitro fertilization/intracytoplasmic sperm injection. *Fertil Steril* 2011;**95**:1650–1654.
- Krassas GE, Poppe K, Glinoe D. Thyroid function and human reproductive health. *Endocr Rev* 2010;**31**:702–755.
- Michalakis KG, Mesen TB, Brayboy LM, Yu B, Richter KS, Levy M, Widra E, Segars JH. Subclinical elevations of thyroid-stimulating hormone and assisted reproductive technology outcomes. *Fertil Steril* 2011;**95**:2634–2637.
- Negro R, Mangieri T, Coppola L, Presicce G, Casavola EC, Gismondi R, Locorotondo G, Caroli P, Pezzarossa A, Dazzi D et al. Levothyroxine treatment in thyroid peroxidase antibody-positive women undergoing assisted reproduction technologies: a prospective study. *Hum Reprod* 2005;**20**:1529–1533.
- Negro R, Formoso G, Mangieri T, Pezzarossa A, Dazzi D, Hassan H. Levothyroxine treatment in euthyroid pregnant women with autoimmune thyroid disease: effects on obstetrical complications. *J Clin Endocrinol Metab* 2006;**91**:2587–2591.
- PEDro scale. 2012. Available at: <http://www.pedro.org.au/english/downloads/pedro-scale/> (12 November 12, date last accessed).
- Poppe K, Glinoe D, Van Steirteghem A, Tournaye H, Devroey P, Schiettecatte J, Velkeniers B. Thyroid dysfunction and autoimmunity in infertile women. *Thyroid* 2002;**12**:997–1001.
- Poppe K, Glinoe D, Tournaye H, Devroey P, Velkeniers B. Impact of the ovarian hyperstimulation syndrome on thyroid function. *Thyroid* 2008;**18**:801–802.
- Reid SM, Middleton P, Cossich MC, Crowther CA. Interventions for clinical and subclinical hypothyroidism in pregnancy. *Cochrane Database Syst Rev* 2010;CD007752.
- Scoccia B, Demir H, Kang Y, Fierro MA, Winston NJ. In vitro fertilization pregnancy rates in levothyroxine-treated women with hypothyroidism compared to women without thyroid dysfunction disorders. *Thyroid* 2012;**22**:631–636.
- Thangaratinam S, Rogozinska E, Jolly K, Glinkowski S, Roseboom T, Tomlinson JW, Kunz R, Mol BW, Coomarasamy A, Khan KS. Effects of interventions in pregnancy on maternal weight and obstetric outcomes: meta-analysis of randomised evidence. *Br Med J* 2012;**344**:e2088. doi: 10.1136/bmj.e2088.
- Toulis KA, Goulis DG, Venetis CA, Kolibianakis EM, Negro R, Tarlatzis BC, Papadimas I. Risk of spontaneous miscarriage in euthyroid women with thyroid autoimmunity undergoing IVF: a meta-analysis. *Eur J Endocrinol* 2010;**162**:643–652.
- Unuane D, Tournaye H, Velkeniers B, Poppe K. Endocrine disorders & female infertility. *Best Pract Res Clin Endocrinol Metab* 2011;**25**:861–873.
- van den Boogaard E, Vissenberg R, Land JA, van Wely M, van der Post JA, Goddijn M, Bisschop PH. Significance of (sub)clinical thyroid dysfunction and thyroid autoimmunity before conception and in early pregnancy: a systematic review. *Hum Reprod Update* 2011;**17**:605–619.
- Vissenberg R, van den Boogaard E, van Wely M, van der Post JA, Fliers E, Bisschop PH, Goddijn M. Treatment of thyroid disorders before conception and in early pregnancy: a systematic review. *Hum Reprod Update* 2012;**18**:360–373.