Thyroxine Suppression Therapy For Benign, Non-Functioning Solitary Thyroid Nodules: A Quality-Effects Meta-Analysis

Altayyeb Yousef, MBBS; Justin Clark, BA, Lib Tech; and Suhail A. R. Doi, MBBS, FRCP, PhD

Corresponding Author:
Suhail A. R. Doi, MBBS, FRCP, PhD
Clinical Epidemiology Unit
School of Population Health
University of Queensland
Brisbane, Australia
Email: sardoi@gmx.net
Fax: +61 7 3365 5599

Received: August 21, 2009
Revised: May 6, 2010
Accepted: May 20, 2010
doi:10.3121/cmr.2010.881
Abstract

Levothyroxine (LT4) suppressive therapy for solitary thyroid nodules is not being popularly advocated presently because its clinical efficacy and safety are currently considered controversial. This meta-analysis aims to address efficacy issues by using rigorous methods to arrive at a pooled estimate. On the basis of the analysis, it is estimated that LT4 therapy is clearly associated with up to a two-fold increase in the chance of nodule reduction. This translates to a number needed to treat (NNT) of 6 or a 50% decrease in the risk of cancer given nodule reduction. Keeping this definition of efficacy in mind and a low incidence of adverse events with low level LT4 suppression, such an intervention might be appropriate in patients selected on the basis of a low risk for adverse effects.

Keywords: Meta-analysis, Levothyroxine suppression, Solitary thyroid nodules, Clinical efficacy, Adverse effects
The natural history of thyroid nodules suggests that a third of benign nodules will spontaneously grow to more than 15% of its initial size within one year and this increases to almost all (89%) at five years if these nodules were not treated somehow.\textsuperscript{1} The majority of growth has been found in those nodules that have more solid component. Although the pathogenesis of such growth is poorly understood, thyroid-stimulating hormone (TSH) is presumed to be necessary if not sufficient, and therefore suppression of TSH secretion might be expected to result in a decrease in nodule size or at least prevent any further enlargement.

Benign solitary nodules detected by physical examination represent two different pathological processes. In many of these patients, the palpated nodule could simply be the dominant nodule of a multinodular goiter while other solitary nodules are probably true solitary adenomas. The latter can be distinguished by ultrasound and nonrandomized, uncontrolled trials have found that patients with solitary nodules had lesser decreases in nodule size in response to thyroxine therapy than patients with nontoxic diffuse or multinodular goiters.\textsuperscript{2} Of the randomized trials that have been published, the majority demonstrated some degree of efficacy of thyroxine therapy for solitary thyroid nodules. Meta-analyses of these trials was first attempted in 1998 and showed a reduction in thyroid solitary nodule volume in 17% of patients treated with levothyroxine and inhibition of nodule growth in another 10%.\textsuperscript{3} A second meta-analysis found that nodules decreased by more than 50% in more patients who received thyroxine therapy, but the treatment response did not reach statistical significance (RR 1.9: 95 percent CI, 0.9 to 3.8).\textsuperscript{4} Also, there have been concerns about subclinical hyperthyroidism in treated patients who may be at increased risk for atrial fibrillation, other cardiac abnormalities, or reduced bone density. These
possibilities, combined with the uncertainty regarding efficacy, has led to recommendations that vary. We therefore decided to re-evaluate studies on this topic with a view to addressing uncertainties regarding efficacy and implications.

**Methods**

*Selection of studies*

A search of the MEDLINE database as well as In-Process Citations was undertaken using PubMed (April 2010; see appendix 1 for search strategy). The PubMed search was then modified for use with Emtree subject headings and used to search the EMBASE database (April 2010; see appendix 2 for search strategy). Reference lists of relevant retrieved studies and review articles were searched for additional studies (only one more abstract retrieved). Grey literature and abstracts were not searched. Results from the searches were combined and after removal of duplication 316 articles were found for possible inclusion.

We then retained studies for detailed review based on a set of inclusion criteria as follows: trials of solitary nodules or multinodular thyroid disease with a prominent single nodule, trials where patients were euthyroid, trials where patients had a benign histology based on fine-needle aspiration cytology, and trials that were placebo controlled. Nineteen studies that passed our initial survey based on meeting all our inclusion criteria were further reviewed, and eight were rejected based on one or more of the following: the presence of a retrospective or cross sectional study design, follow-up data not available or not extractable, no outcome in either group, and finally the failure to use ultrasound to decide the outcome of the hormonal therapy on the size of
the thyroid nodule under investigation (table 1A) leaving us with 11 studies (table 1B) that were pooled for this meta-analysis.

Quality assessment and statistical analysis

Quality was assessed using a study-specific modification of the scoring system published by Doi and Thalib.\textsuperscript{5} Table 2 shows details of the quality scoring criteria. The dichotomous variable ‘presence of nodule reduction’ was determined by the various criteria used in individual studies as detailed in table 1. Table 1 also shows, for each study, the proportion of thyroxine and placebo subjects who achieved the defined goal of therapy in that study. For each study the relative risk (RR) was used as a measure of the relation between treatment status and risk of nodule reduction. A value of 1 indicates equivalence and a value more than 1 indicates a higher risk of reduction for the thyroxine treated groups. For graphical purposes the RR was plotted on a log scale as presentation is clearer because the confidence intervals become symmetrical about the point estimates of the RR. Pooled results were calculated via both the fixed-effect and quality-effect models. This analysis was done using MIX v 1.7.\textsuperscript{6} To assess the robustness of our meta-analysis, given the possibility of publication bias, we computed the fail-safe N, which is the number of studies with negative findings that would need to be combined with the studies reviewed to lead to a non-significant result. The larger the fail-safe N, the less likely it is that unpublished studies or future studies would overturn our result. Publication bias was also assessed using funnel plot asymmetry. We considered the funnel plot to be asymmetrical if the slope of Macaskill’s regression line deviated from zero with a $P$ value of less than 0.10.
Results

Quantitative synthesis

Of 19 studies identified, data from 11 were included\textsuperscript{3,7-16} and eight studies were excluded. Of these eight, we sent a request for data to the authors of three studies\textsuperscript{17-19} as the data in the papers were insufficient to decide how many people failed in either one or both treatment groups. As of the time of submission of this paper, the authors had not responded to our requests for the missing data. Two others excluded were retrospective or cross-sectional\textsuperscript{20,21} while two more\textsuperscript{22,23} were excluded for reasons given in table 1A. The last (Uzunkoy et al. Levothyroxine suppressive therapy for solitary thyroid nodules. Poster Session. 75th Annual Meeting of the American Thyroid Association. 2003) was excluded because it remained unpublished.

There were a total of 743 patients: 417 patients were given thyroxine and 326 patients given placebo. Of the 11 studies, eight were randomized controlled trials (six single center studies and two multicenter studies) and three were cohort studies. Because of concern about combining data from studies with markedly different designs, separate analyses were done for each group as well as a pooled analysis. The final pool of eight randomized and three cohort studies are shown in table 1. The trials were deemed heterogeneous (because $\tau^2 = 0.06$ even though $Q = 15.9$; $P = 0.1$), and were not deemed combinable in one stratum using a fixed effects model, and thus we used the quality effects model to pool data.\textsuperscript{5,24} We did not use a random effects model because we do not think it has any clear interpretation in meta-analysis.\textsuperscript{25}
The summary thyroxine to placebo group RR of reduction after the end of follow-up was 2.95 (95% confidence interval [CI], 1.2-7.2) for the cohort studies, 1.34 (95% confidence interval [CI], 0.92 – 1.95) for the single center randomized trials, and 1.75 (95% confidence interval [CI], 1.3-2.4) for the multi-center trials using the quality effects model. The summary and individual RRs and 95% CI's are depicted individually in figure 1A and cumulatively in figure 1B in increasing order of the risk of reduction with placebo therapy. The pooled RR for all studies was 1.68 (95% confidence interval [CI], 1.3-2.1). This indicates that thyroxine treated patients are up to twice as likely as placebo treated patients to have nodule reduction at the end of follow up (minimum follow up 6 months).

Sensitivity analysis and publication bias

To determine how sensitive our findings were to changes in the data included in our analysis, we analyzed the effect of sub-grouping of the various studies on the pooled results. The summary treatment to control group RR for the 11 studies was determined after being re-grouped for sample size (<40 patients and >40 patients in the thyroxine group), risk of success in placebo group (<20% and ≥20%), year of publication (in or prior to 1995 and after 1995), and model used for the pooled analysis (fixed, random and quality effects model). Results are depicted in table 4, and what is striking is that both the extent of TSH lowering as well as a lower risk of success in the placebo groups are key determinants of the relative success of thyroxine suppression. One-way sensitivity analyses with each trial individually removed suggests that no trial individually altered the pooled results significantly (figure 2).
Analysis of the effect of potential unpublished or missed negative result studies was undertaken to evaluate the robustness of our meta-analysis. This sensitivity analysis to publication bias using the fail-safe N technique reveals that 48 studies of size equal to the average of those in our meta-analysis and showing no benefit for thyroxine therapy would have to have been missed to change the statistical significance of our results. Also, the value for the slope on Macaskill regression for these studies was –0.01 (non-weighted) but this not significantly different from zero \( (P = 0.47) \) suggesting no publication bias.

**Failure rates after placebo therapy for nodule reduction**

The average reduction rate for the placebo therapy was 23% (range 2% – 60%). This wide variation in nodule reduction rates is related to differences in the definition of nodule reduction and the variation in the extent of nodule size and content.

**Implications of the quantitative synthesis**

From the quantitative synthesis above, the pooled treatment to placebo RR is 1.7 (rounded off) and therefore the increase in relative risk of reduction using thyroxine therapy is 70%. Given that the average chance of nodule reduction with placebo is about 23% (see above), then the absolute increase in risk with thyroxine would be 16% given the relative risk increase of about 70%. The numbers of patients needed to be treated \((NNT)^{26}\) with thyroxine to lead to one more nodule reduction would then be about six. This means that one more success will occur for every six patients given thyroxine compared to no treatment.
Discussion

Thyroid-hormone-suppressive therapy with levothyroxine (LT4) is still being used by endocrinologists for the treatment of solitary thyroid nodules\textsuperscript{27} despite the fact that such use has been variably recommended in currently accepted treatment guidelines.\textsuperscript{28,29} This is probably because it is thought that LT4-induced suppression of TSH secretion shrinks thyroid nodules by preventing TSH’s growth-promoting effect on thyroid cells.\textsuperscript{30} In support of this line of reasoning, we were indeed able to demonstrate a decrease in nodule volume in one more patient for every six treated with LT4 compared to placebo. However, this was a surrogate endpoint that does not necessarily correlate with the patient’s clinical outcome, although there is reason to believe that nodule shrinkage is a beneficial aspect of treatment (see below). Although five previous meta-analyses\textsuperscript{3,4,31-33} have been published, this is the first to carefully assess the quality of included studies, as well as incorporate that in the inverse-variance adjustment of study weights. Also, we strictly restricted assessment to that based on ultrasonography of the solitary nodule. Overall, our analysis concurred with previous estimates of an approximately two-fold increase in nodule reduction with LT4 suppression. One of these meta-analyses\textsuperscript{33} also assessed causality using the Hill criteria, which suggest that a two-fold increased chance of reducing thyroid nodule by $>\text{50\%}$ seems plausible to recommend using LT4 for treatment.

Many of the trials we combined in this analysis were non-randomized and had a wide variation in the number of patients, target levels of TSH suppression and treatment duration. There was also considerable variation in the degree of volume reduction used to define therapeutic efficacy. Also, eight of the 19 selected studies were excluded, which seems to be a significant number
(and some were relatively large studies). If they were expected to be substantially different from the 11 included studies this could bias the overall analysis. However, there was no reason to suspect this, and there was nothing we could do to include them. Despite these limitations, we can state that, on average, the effects of LT4 lead to an almost two-fold greater chance (compared to placebo) of LT4-treated patients to show nodule shrinkage. Sensitivity analysis suggests that this effect is consistent across all patient groups and the varying RR’s are a result of the propensity of various patient populations studied in this meta-analysis to spontaneously have nodule reduction. In other words, those groups of patients whose nodules do not shrink spontaneously with placebo therapy are those where LT4 has the maximum efficacy. In the latter group, iodine supplementation can possibly do the same as thyroxine and could be a source of confounding in our data. This threshold seems to be at 20%, suggesting that LT4 suppression leads to at least a further 20% incidence of nodule reduction in these treated populations. The concern with prolonged thyroid hormone-suppressive therapy is that it induces a state of subclinical hyperthyroidism. This condition can be the source of unpleasant symptoms as well as more serious adverse effects involving the cardiovascular system, but this is most likely in those patients with evidence of mild hyperthyroidism and adrenergic overactivity. Another concern is the effect of LT4 suppression on bone metabolism, but this seems to be a problem predominantly among postmenopausal women but not among premenopausal women. If treatment is aimed at reduced but not totally suppressed TSH values, there is less effect on the skeleton. Overall, the consequences of low level subclinical thyroid disease (serum TSH 0.1-0.45 mU/L) seem to be minimal, and there is insufficient evidence to support complications, except in pregnant women, women older than 60 years, and others with cardiovascular morbidity. By contrast, there was fair-to-good evidence for increased cardiac manifestations among patients with serum
TSH levels below 0.10 mIU/l, although the evidence for bone alterations in these patients was less impressive.40 Overall, evidence is conflicting, and while one meta-analysis concluded that evidence of an association between heart disease events and mortality in subclinical thyrotoxicosis was weak,39 another meta-analysis concluded that untreated endogenous subclinical thyrotoxicosis is more harmful in patients with co-morbidities such as cardiac disease and diabetes mellitus, and in patients recovering from stroke.38

Even though a greater degree of LT4 suppression was associated with a better outcome, this was not as important as the level of control group risk, suggesting that lower level TSH suppression (0.1 to 0.3 mU/L) might be as useful as more suppressed therapy, and this may be a solution to concerns about the risk of subclinical thyrotoxicosis. Indeed, it has been suggested by one of the studies excluded from this meta-analysis (due to insufficient data) that LT4 should be given at doses that reduce rather than abolish TSH secretion.17 In this placebo-controlled, randomized, crossover trial, low-level and high-level suppression of TSH (target levels 0.40–0.60 mIU/L versus <0.01 mIU/l) proved to be equally effective in reducing thyroid nodule size.17

This meta-analysis result becomes important if we accept the hypothesis that cancers are less likely to regress than benign nodules in response to normal neuroendocrine regulatory pathways. In two studies of thyroid cancer, 13% to 15% (11/83 and 4/26 respectively) of malignant nodules were found to regress with thyroxine.41,42 This is even below the response rate of 23% in our placebo group, and far below the 70% increase in response rates with thyroxine, thus lending credence to this hypothesis. If we assume that a benign nodule has a risk of response to thyroxine of 39% (23*1.7) and a malignant nodule, a response to thyroxine of 15%, the negative likelihood
ratio for suppression is $15/39 = 0.4$. Assuming a baseline risk of malignancy of 5%, a nodule reducing in size after thyroxine suppression would be expected to have a reduction in chance of harbouring a cancer by at least half its original risk.

In summary, LT4 can produce significant volume reductions in benign solitary thyroid nodules, but its use is inappropriate in certain groups of patients, such as those over 60 years of age, elderly patients and postmenopausal women. In younger patients without co-morbidity, low level LT4 suppression will reduce nodule size in at least one more out of every six treated patients with benign nodules, and in these patients the chance of malignancy would be decreased by half.
References


24. Doi SA, Thalib L. An alternative quality adjustor for the quality effects model for meta-analysis. Epidemiology 2009;20:314


40. Surks MI, Ortiz E, Daniels GH, Sawin CT, Col NF, Cobin RH, Franklyn JA, Hershman JM, Burman KD, Denke MA, Gorman C, Cooper RS, Weissman NJ. Subclinical thyroid disease: scientific review and guidelines for diagnosis and management. JAMA 2004;291:228-238.


Author Affiliations

Altayyeb Yousef, MBBS*; Justin Clark, BA, Lib Tech†, and Suhail A. R. Doi, MBBS, FRCP, PhD‡

*Department of Medicine, Mubarak Al-Kabeer Teaching Hospital, Kuwait
†Herston Health Sciences Library, University of Queensland, Brisbane, Australia
‡Clinical Epidemiology Unit, School of Population Health, University of Queensland, Brisbane, Australia
<table>
<thead>
<tr>
<th>Study Name</th>
<th>No. of patients involved (T4 / Placebo)</th>
<th>Duration of treatment/ followup</th>
<th>Criteria of non-reduction</th>
<th>Quality score</th>
<th>Reason for exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Costante 2004</td>
<td>43/38</td>
<td>1 year</td>
<td>any non-reduction</td>
<td>0.5</td>
<td>Retrospective study</td>
</tr>
<tr>
<td>2 Baldini 2002</td>
<td>43/46</td>
<td>1 year</td>
<td>any non-reduction</td>
<td>0.5</td>
<td>Cross-sectional study</td>
</tr>
<tr>
<td>3 Papini 2007</td>
<td>21/21</td>
<td>1 year</td>
<td>Nodule is not reduced or reduced but to &lt; 50%</td>
<td>0.83</td>
<td>There is no outcome in the placebo nor in treated group</td>
</tr>
<tr>
<td>4 Tsai 2006</td>
<td>30/30</td>
<td>6 months</td>
<td>Nodule is not reduced or reduced but to &lt; 50%</td>
<td>0.75</td>
<td>No data</td>
</tr>
<tr>
<td>5 Uzunkoy 2003</td>
<td>50/50</td>
<td>12 months</td>
<td>Nodule is not reduced or reduced but to &lt; 50%</td>
<td>-</td>
<td>Study was reported in a conference. No available data in publication.</td>
</tr>
<tr>
<td>6 Koc 2002</td>
<td>20/20</td>
<td>1 year</td>
<td>Nodule is not reduced or reduced but to &lt; 50%</td>
<td>0.67</td>
<td>No extractable data</td>
</tr>
<tr>
<td>7 Berghout 1990</td>
<td>26/26</td>
<td>9 months</td>
<td>Nodule is not reduced or reduced but to &lt; 13%</td>
<td>0.75</td>
<td>Data regarding placebo is not complete Solitary nodules sample are small (3 patients)</td>
</tr>
<tr>
<td>8 Cheung 1989</td>
<td>37/37</td>
<td>36 months</td>
<td>Nodule is not reduced or reduced but to &lt; 50%</td>
<td>0.5</td>
<td>The thyroid measurement is palpation based</td>
</tr>
</tbody>
</table>
Table 1B. Included studies.

<table>
<thead>
<tr>
<th>Study Name</th>
<th>No. of patients (T4 / Placebo)</th>
<th>Average Dose of thyroxine (μg)</th>
<th>Average TSH level at the end of trial for T4/Placebo mIU/L</th>
<th>Duration of treatment / followup</th>
<th>Criteria of failure of reduction after treatment</th>
<th>Non-reduction T4/Placebo</th>
<th>Q</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Papini 1998</td>
<td>42/41</td>
<td>2 ug /kg /day</td>
<td>0.11 +- 0.06 / 1.59 +- 0.43</td>
<td>60 months</td>
<td>Nodule is not reduced or reduced but to &lt;11.7%</td>
<td>30/22</td>
<td>0.75</td>
</tr>
<tr>
<td>2 Papini 1993</td>
<td>51/50</td>
<td>2 ug /kg /day</td>
<td>0.06 +- 0.06/ 1.08 +- 0.52</td>
<td>12 months</td>
<td>Failure of Any reduction</td>
<td>28/37</td>
<td>0.75</td>
</tr>
<tr>
<td>3 Gharib 1987</td>
<td>28/25</td>
<td>3 ug /kg /day</td>
<td>0.5 +- 0.4 / 1.6 +- 0.7</td>
<td>6 months</td>
<td>Failure of any reduction</td>
<td>14/10</td>
<td>0.75</td>
</tr>
<tr>
<td>4 Zelmanovitz 1998</td>
<td>21/24</td>
<td>2.7 +- 0.3 / ug/kg</td>
<td>0.24 +- 0.33 / 1.17 +- 0.65</td>
<td>1 year</td>
<td>Nodule is not reduced or reduced but to &lt;50%</td>
<td>15/22</td>
<td>0.83</td>
</tr>
<tr>
<td>5 Sakalauskiene 2002</td>
<td>37/25</td>
<td>58.8 +- 30.97 (fixed dose 200 ug/day) average dose 2.48+- 0.9 ug/kg/day</td>
<td>1.28 +- 0.71 / 1.52 +- 0.74</td>
<td>6 month</td>
<td>Failure of any reduction</td>
<td>24/21</td>
<td>0.33</td>
</tr>
<tr>
<td>6 Lima 1997</td>
<td>54/20</td>
<td>2.24 +- 0.45 ug /kg /day</td>
<td>&lt;0.1 /</td>
<td>12 months</td>
<td>Nodule is not reduced or reduced but to &lt;20%</td>
<td>23/16</td>
<td>0.67</td>
</tr>
<tr>
<td>7 Wemeau 2002</td>
<td>64/59</td>
<td>2.24 +- 0.45 ug /kg /day</td>
<td>Lowered by 0.73 / increased by 0.05</td>
<td>18 months</td>
<td>Nodule is not reduced or reduced but to &lt;20%</td>
<td>30/41</td>
<td>0.75</td>
</tr>
<tr>
<td>8 Reverter 1992</td>
<td>20/20</td>
<td>2.82 +- 0.6 ug /Kg /day</td>
<td>&lt; 0.1 / 1.1 +- 0.8</td>
<td>1 year</td>
<td>Nodule is not reduced or reduced but to &lt;20%</td>
<td>16/17</td>
<td>0.83</td>
</tr>
<tr>
<td>9 La Rosa</td>
<td>23/22</td>
<td>1.94 +- 0.16 ug/kg</td>
<td>0.1 / 1.2</td>
<td>1 year</td>
<td>Nodule is not reduced or reduced but to &lt;50%</td>
<td>14/22</td>
<td>0.92</td>
</tr>
<tr>
<td>10 Mainini</td>
<td>45/10</td>
<td>1.7ug/kg</td>
<td>&lt;0.1 / 1.7+- 0.9</td>
<td>2 years</td>
<td>Nodule is not reduced or reduced but to &lt;50 %</td>
<td>37/10</td>
<td>0.67</td>
</tr>
<tr>
<td>11 Larijani 1999</td>
<td>32/30</td>
<td>1.5 – 2.0 ug / Kg / day</td>
<td>0.17 +- 0.2 / no record of control TSH at the end of trial, N.B this is TSH after induction with TRH</td>
<td>12 months</td>
<td>Failure of any reduction</td>
<td>17/18</td>
<td>0.92</td>
</tr>
</tbody>
</table>
### Table 2. Quality score criteria

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Description of the score</th>
<th>Points range</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Method of randomization</td>
<td>Not reported or method reported as non randomized or multicenter = 0  Authors stated that the study was randomized but details of randomization procedure are vague or missing = 1  Randomized single center study with a clear description of randomization scores 2  Allocating patients was not concealed and was done by someone who knew patients’ history or there was no comment on allocation in the study scores 0</td>
<td>0-2</td>
</tr>
<tr>
<td>2 Concealed treatment allocation:</td>
<td>The allocation was not concealed but the allocating individual had no knowledge about the patient history scores 0.5  Totally Concealed allocation scores 1</td>
<td>0-1</td>
</tr>
<tr>
<td>Prognostic indicators at baseline: Size of nodules Constituent of the nodule</td>
<td>To score 1: There is ≤ 3 of these prognostic factors mismatched at baseline. Score of 0 is where: &gt; 3 mismatched criteria at the baseline of study</td>
<td>0 – 1</td>
</tr>
<tr>
<td>3 Age of patient Dose of thyroxine Duration of follow up Eligibility criteria: Solitary nodules Euthyroid patient Benign histology based on FNAC No comorbid conditions or medication that could interfere with thyroxine No prior thyroid surgery No exposure to irradiation of head and neck Prior thyroxin treatment</td>
<td>outcome was not present = 1</td>
<td>0-1</td>
</tr>
<tr>
<td>4</td>
<td>No blinding scores 0  single blinded scores 0.5  double blinded scores 1</td>
<td>0-1</td>
</tr>
<tr>
<td>5 The outcome of the study was present at the onset of the study</td>
<td>To score 0: &lt; 3 criteria should be present  To score 1: &gt;4 of these criteria should be present</td>
<td>0-1</td>
</tr>
<tr>
<td>6 Blinding:</td>
<td>Enough assessment is defined as &gt;6 months treatment of T4  &lt; 6 months scores 0  &gt;6 months scores 1</td>
<td>0-1</td>
</tr>
<tr>
<td>7 Protocol deviations</td>
<td>Intervention and outcomes are clearly defined. In which outcome: 1) TSH was adequately suppressed (TSH &lt;0.3) and 2) clear method of assessing change in nodule size clearly stated</td>
<td>0-1</td>
</tr>
<tr>
<td>8 Was the timing of outcome assessment enough:</td>
<td>Yes both are achieved= 1  No = 0</td>
<td>0-1</td>
</tr>
<tr>
<td>9</td>
<td>Yes = 1  No=0</td>
<td>0-1</td>
</tr>
<tr>
<td>10</td>
<td>Yes = 1  No=0</td>
<td>0-1</td>
</tr>
<tr>
<td>11</td>
<td>Maximum total points = 12  Quality score = (sum of points) / 12 (Range 0 to 1)</td>
<td>0-1</td>
</tr>
</tbody>
</table>
Table 3. Sensitivity analysis

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Pooled RR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thyroxine group sample size</td>
<td></td>
</tr>
<tr>
<td>≤ 40 subjects</td>
<td>1.4 (0.98 - 2)</td>
</tr>
<tr>
<td>&gt; 40 subjects</td>
<td>1.96 (1.4 – 2.7)</td>
</tr>
<tr>
<td>Risk of success in the placebo group</td>
<td></td>
</tr>
<tr>
<td>&lt; 20%</td>
<td>3.2 (1.4 - 7.1)</td>
</tr>
<tr>
<td>≥ 20%</td>
<td>1.5 (1.2-1.8)</td>
</tr>
<tr>
<td>Year of publication</td>
<td></td>
</tr>
<tr>
<td>≤ 1995</td>
<td>1.6 (1 – 2.4)</td>
</tr>
<tr>
<td>&gt; 1995</td>
<td>1.8 (1.3 – 2.4)</td>
</tr>
<tr>
<td>Meta-analysis model</td>
<td></td>
</tr>
<tr>
<td>Quality effects</td>
<td>1.7 (1.3 - 2.1)</td>
</tr>
<tr>
<td>Random effects</td>
<td>1.6 (1.2 - 2.1)</td>
</tr>
<tr>
<td>Fixed effects</td>
<td>1.7 (1.4 - 2.1)</td>
</tr>
<tr>
<td>Degree of TSH suppression in treatment groups</td>
<td></td>
</tr>
<tr>
<td>(excludes Lima &amp; Mainini due to lack of data)</td>
<td></td>
</tr>
<tr>
<td>≤ 0.1 mU/L</td>
<td>2.1 (1.4 – 3.2)</td>
</tr>
<tr>
<td>&gt; 0.1 mU/L</td>
<td>1.3 (1 – 1.8)</td>
</tr>
</tbody>
</table>
Figures

Figure 1 (A): Forest plot of the 11 studies included in this meta-analysis. Size of the boxes are proportional to the relative weights in each study.
Figure 1 (B): Cumulative forest plot of the same studies as in figure 1a in order of year of publication. A significant effect is evident from the year 1995.
**Figure 2:** Exclusion sensitivity plot. There was no significant effect of the exclusion of any single trial suggesting results are robust.
Appendix 1. PubMed search

#45: Search (((#34) AND #41) AND #42) AND #43) AND #44

#44: Search Randomized OR Randomised OR Placebo OR Randomly OR Random OR Trial OR Groups OR Group OR Blinded

#43: Search Growth OR Increase OR Decrease OR Size OR Volume

#42: Search TSH-suppressive OR Suppression OR Suppress OR Suppressive

#41: Search (((#35) OR #36) OR #37) OR #40

#40: Search (#38) AND #39

#39: Search Nodule OR Nodules OR Nodular

#38: Search “Thyroid Gland”[MeSH] OR Thyroid

#37: Search Goiter

#36: Search “Goiter”[MeSH]

#35: Search "Thyroid Nodule"[Mesh]

#34: Search (((#29) OR #30) OR #31) OR #32) OR #33

#33: Search Dexnon OR Eferox OR Eltroxin OR Thevier OR Eltroxine OR Euthyrox OR Eutirox OR LevoT OR “Levo T” OR Levo-T OR Levothyroid OR Levothyroxin

#32: Search Synthroid OR Synthrox OR Thyrax OR Tiroidine OR “Tiroxina Leo” OR Unithroid OR Berlthyrox

#31: Search Levothroid OR Levoxine OR Levoxyl OR Lévothyrox OR Novothyral OR Novothyrox OR Oroxine

#30: Search Thyroxine OR Thyroxin OR “Thyroid Hormone” OR L-thyroxine OR Tetraiodothyronine OR Levothyroxine

#29: Search "Thyroxine"[MeSH]
Appendix 2. Embase search

#15: #9 AND #10 AND #11 AND #12 AND #13 AND [embase]/lim
#14: #9 AND #10 AND #11 AND #12 AND #13
#13: #5 OR #8
#12: #1 OR #2 OR #3 OR #4
#11: randomized OR randomised OR 'placebo'/exp OR placebo OR randomly OR random OR trial OR groups OR group OR blinded
#10: 'growth'/exp OR growth OR increase OR decrease OR size OR volume
#9: 'tsh suppressive' OR suppression OR suppress OR suppressive
#8: #6 AND #7
#7: nodule OR nodules OR nodular
#6: 'thyroid gland'/exp OR 'thyroid gland' OR 'thyroid'/exp OR thyroid
#5: 'thyroid nodule'/exp OR 'thyroid nodule' OR 'goiter'/exp OR goiter
#4: dexnon OR 'eferox'/exp OR eferox OR 'eltr oxin'/exp OR eltroxin OR 'thevier'/exp OR thevier OR eltroxine OR 'euthyrox'/exp OR euthyrox OR 'eutirox'/exp OR eutirox OR levot OR 'levo t' OR 'levothyroid'/exp OR levothyroid OR levothyroxin
#3: 'synthroid'/exp OR synthroid OR synthrox OR thyrax OR tiroxina leo' OR unithroid OR berlthyrox
#2: 'levoxine'/exp OR levoxine OR 'levoxyl'/exp OR levoxyl OR lévothyrox OR 'novothyral'/exp OR novothyral OR novothyrox OR 'oroxine'/exp OR oroxine
#1: 'thyroxine'/exp OR thyroxine OR 'thyroxin'/exp OR thyroxin OR 'thyroid hormone'/exp OR 'thyroid hormone' OR 'l thyroxine'/exp OR 'l thyroxine' OR 'tetraiodothyronine'/exp OR tetraiodothyronine OR 'levothyroxine'/exp OR levothyroxine OR 'levothroid'/exp OR levothroid