National Institute for Health and Care Excellence

Draft for Consultation

Thyroid disease: assessment and management

[G] Management of subclinical hypothyroidism

NICE guideline

Intervention evidence review underpinning recommendations 1.5.1 to 1.5.9 in the guideline

June 2019

Draft for Consultation

This evidence review was developed by the National Guideline Centre

1

Disclaimer

The recommendations in this guideline represent the view of NICE, arrived at after careful consideration of the evidence available. When exercising their judgement, professionals are expected to take this guideline fully into account, alongside the individual needs, preferences and values of their patients or service users. The recommendations in this guideline are not mandatory and the guideline does not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of the individual patient, in consultation with the patient and, where appropriate, their carer or guardian.

Local commissioners and providers have a responsibility to enable the guideline to be applied when individual health professionals and their patients or service users wish to use it. They should do so in the context of local and national priorities for funding and developing services, and in light of their duties to have due regard to the need to eliminate unlawful discrimination, to advance equality of opportunity and to reduce health inequalities. Nothing in this guideline should be interpreted in a way that would be inconsistent with compliance with those duties.

NICE guidelines cover health and care in England. Decisions on how they apply in other UK countries are made by ministers in the Welsh Government, Scottish Government, and Northern Ireland Executive. All NICE guidance is subject to regular review and may be updated or withdrawn.

ISBN

Contents

1

1 Management of subclinical hypothyroidism

1.1 Review question: What is the clinical and cost-effectiveness of treating subclinical hypothyroidism?

1.2 Introduction

Subclinical hypothyroidism (SCH) is a biochemical state in which the serum TSH is elevated above the reference range whilst the concentrations of circulating free thyroid hormones (FT4, FT3) are within the reference range for the population. It is more common in women and becomes increasingly prevalent with age, such that around 5% of people over 70 years of age, and 10% of people over 80 may manifest SCH.

In some people SCH may be a transient phenomenon reflecting non-thyroidal illness or drug effects, but in others it may be an early manifestation of a disease process such as Hashimoto's thyroiditis- in this situation the biochemical picture represents a state of compensated or mild hypothyroidism in which the circulating thyroid hormones remain within the reference range owing to increased thyroid gland stimulation by TSH.

As serum free thyroid hormones are within the normal range for the population in SCH there is uncertainty as to whether people benefit from increasing their circulating thyroid hormones with replacement therapy, Symptoms of hypothyroidism are non-specific and common in the euthyroid population, meaning that clinicians cannot be confident they have treated symptoms caused by hypothyroidism in someone with SCH. This leaves open several questions about the optimal management of people with persistent SCH.

1.3 PICO table

For full details see the review protocol in Appendix A:.

Table 1: PICO characteristics of review question

Population	People diagnosed with subclinical hypothyroidism (TSH greater than upper limit of context specific normal range, T3/T4 within normal range)
Interventions	T3 T4-initiation at high dose T4-initiation via gradual titration Combination of T3 & T4 Natural thyroid extract (mammalian only) Iodine supplementation Selenium supplementation Placebo
Comparisons	Any above vs any other, in isolation or combination
Outcomes	**Critical** • Mortality (dichotomous, ≥1 year) • Quality of life (continuous) **Important** • Cardiovascular morbidity-ischemic heart disease, heart failure (dichotomous) • Arrhythmias (dichotomous) • Osteoporosis (dichotomous) • Impaired cognitive function (dichotomous)

	• Depression (dichotomous) • Patient/family/carer experience of care (continuous) • Healthcare contacts (rates/dichotomous) • Symptom scores (continuous) • Growth (continuous) • TSH suppression (dichotomous)
Study design	• RCTs only • Blinded comparisons prioritised, non-blinded comparisons only considered if blinded unavailable on an intervention by intervention basis • Minimum treatment duration of 3 months

1.4 Clinical evidence

1.4.1 Included studies

Six RCTs were included in the review;[19, 24, 28, 34, 35, 40] these are summarised in Table 2 below. Evidence from these studies is summarised in the clinical evidence summary below (Table 3). One Cochrane review was identified in this area [45], the studies included in this publication were checked against the protocol and were not included in this review.

All six RCTs compared T4 with placebo.

No relevant clinical trials comparing T3, natural thyroid extract, and iodine or selenium supplementation with any other intervention or placebo were identified.

Five included studies were in the adult (18-65) age stratum, whereas one study was in the older adult (>65) age stratum. [40] The majority of participants were female in four studies; [28, 34, 35, 40] whereas, the remaining two studies were conducted exclusively on female participants. [19, 24] Four RCTs were conducted on a treatment naïve population. In one RCT participants having received thyroid medication in the past 3 months were excluded [24] whereas whether participants had received previous treatment was not specified in one RCT [35] The primary cause of subclinical hypothyroidism was autoimmune thyroiditis in at least four studies [19, 24, 34, 35]; whereas this was not reported in the remaining two studies.[28, 40]

The follow-up period of the included studies was from 3 to 12 months.

See also the study selection flow chart in Appendix C:, study evidence tables in Appendix D:, forest plots in Appendix E: and GRADE tables in Appendix F:.

1.4.2 Excluded studies

See the excluded studies list in Appendix J:.

1.4.3 Summary of clinical studies included in the evidence review

Table 2: Summary of studies included in the evidence review

Study	Intervention and comparison	Population	Outcomes	Comments
Kong 2002[19]	T4: 50-100 µg daily (depending on TSH >5 µU/mL) , n= 23 Placebo: ascorbic acid, 5 mg daily, n=17	Women (T4 mean age 53 SD 3, placebo mean age 45, SD 4) TSH 5-10 µU/mL, T4 0.8-16ng/dL; baseline TSH (µU/mL): mean (SD) T4 8 (1.5); Placebo 7.3 (1.6) Symptom status: All except for two patients had self-reported symptoms suggestive of hypothyroidism ;reported to have mild subclinical hypothyroidism Treatment naïve UK	Symptom scores 6 months	80 % autoimmune thyroiditis Parallel design Post-treatment TSH mean (SD) (µU/mL): T4 -4.6 (2.3); placebo -1.7 (2)
Meier 2001 [24]	T4: 25, 50, 75, 100 or 125 µg daily, adapted every 6 weeks to achieve euthyroid TSH (0.1-4.0 mIU / liter) , mean dose at end of study: 85.5 SD 4.3, n=31 Placebo: similar dose and adjustments , n=32	Women (T4 mean age 57.1 SD 10.34, placebo mean age 57.1 SD 10.91) TSH > 5mIU/L, exaggerated TSH response of more than 35 mIU/L after oral TRH stimulation; baseline TSH range 5.0-50 mIU/L, mean (SD) TSH (mIU/L): T4 12.8	Symptom scores 12 months	50% autoimmune thyroiditis, 33% RAI or SUR for Graves' disease Parallel design TSH reference range 0.1-4.0 mIU/L Post treatment TSH mean (SD) (mIU/liter): T4 3.1 (1.67);

Study	Intervention and comparison	Population	Outcomes	Comments
		(7.79); Placebo 10.7 (5.09) Symptom status: not reported Treatment naive Switzerland		Placebo:9.9 (3.94)
Najafi 2015 [28]	T4: 100 µg daily, n=30 Placebo, n=30	Adults (mean age: 34, SD: 10) 85% female TSH > 4.5 mIU/L, normal FT4 (0.8-2 ng/dl), positive anti-TPO-Ab; baseline TSH (mIU/L)mean (SD): T4 8.29 (4.9) Placebo 8.12 (3.12) Symptom status: symptomatic (most prevalent symptoms: weigh gain, fatigue, muscle cramp, irregular menstruation, limb numbness; no significant improvement post treatment) Treatment naïve Iran	Depression TSH suppression 3 months	Parallel design Post treatment TSH mean (SD) mIU/L: T4 2.01 (1.34); Placebo 7.82 (5.17)

Study	Intervention and comparison	Population	Outcomes	Comments
Razvi 2007[34]	T4: 100 µg daily, n=100 Placebo, n=100	Adults (mean age: 53.8, SD: 12) 81% female TSH> 4mIU/L, FT4 0.7-1.9 ng/dl; baseline TSH (mIU/L) median (range): 5.3 (3.7-15.8); T4 first 5.4 (3.8-15.8); Placebo first 5.3 (3.7-13.9) Symptom status: not reported Treatment naïve United Kingdom	Hypothyroid-dependent quality of life TSH suppression (<0.4 mIU/L) 3 months	51% Autoimmune thyroiditis Cross-over design TSH reference range 0.4-4.0 mIU/L Post treatment TSH median (range) mIU/liter : T4 0.5 (0.01-12); Placebo 5.2 (0.9-63.4)
Reuters 2012 [35]	T4: 25 mcg, 50 mcg or 75 mcg depending on stratification by TSH levels (adjusted at 2 months), n=35 Placebo: dose adjustment by TSH levels, n=36	Adults (mean age: 50.01, SD:10.89) TSH > 4 µU/mL, normal FT4 (0.9-1.8 ng/dL); baseline TSH (µU/mL) mean (SD) : T4 7.3 (2.3); Placebo 7.6 (2.7) No information provided on previous treatment Symptom status not reported 87.3% female	Quality of life Depression 6 months	68 % Autoimmune thyroiditis Parallel design TSH Reference range 0.4-4.0 mIU/L Post-treatment TSH not reported.

Study	Intervention and comparison	Population	Outcomes	Comments
		Brazil		
Stott 2017 [40]	T4: 50 µg daily (or 25 if body weight <50 kg or known coronary heart disease), adjusted according to TSH (0.40 - 4.59 mIU/L), n=368 Placebo: with mock dose adjustments, n=369	Older adults (mean age: 74.4, SD: 6.3) 53.7% female TSH 4.60 to 19.99 mIU/L, FT4 within reference range; baseline TSH (mIU/L) mean (SD) overall 6.40 (2.01), T4 6.41 (2.01), Placebo 6.38 (2.01) Symptom status: 27% asymptomatic (based on hypothyroid symptom scale) Mean baseline hypothyroid symptoms score: T4 17.5 (18.8); Placebo 16.9 (17.9); Mean baseline tiredness score: T4 25.9 (20.6); Placebo 25.5 (20.3) Treatment naïve United Kingdom, Netherlands, Switzerland	Health related quality of life Symptom scores 12 months	Multicentre Parallel design Post treatment TSH mean (SD) (mIU/liter): T4 3.63 (2.11); Placebo 5.48 (2.48) TSH at extended follow up visit mean (SD) (mIU/liter); T4 3.47 (2.08); Placebo 5.28 (2.50)

See Appendix D: for full evidence tables.

1.4.4 Quality assessment of clinical studies included in the evidence review

Table 3: Clinical evidence summary: T4 vs Placebo in adults

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo	Risk difference with T4 (95% CI)
QoL: Hypothyroid-dependent T-QoL. Scale from: -3 to 1. Better indicated by higher values.	100 (1 study) 12 months	⊕⊕⊕⊖ MODERATE2 due to risk of bias		The mean qol: hypothyroid-dependent in the control groups was -1.2	The mean qol: hypothyroid-dependent in the intervention groups was 0.1 higher (0.16 lower to 0.36 higher)
QoL: General health SF-36. Scale from: 0 to 100. Better indicated by higher values.	57 (1 study) 6 months	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		The mean qol: general health in the control groups was 7.4	The mean qol: general health in the intervention groups was 0.9 lower (9.08 lower to 7.28 higher)
QoL: Physical functioning SF-36. Scale from: 0 to 100. Better indicated by higher values.	57 (1 study) 6 months	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		The mean qol: physical functioning in the control groups was 1.9	The mean qol: physical functioning in the intervention groups was 1.8 higher (8.27 lower to 11.87 higher)
QoL: Role-physical SF-36. Scale from: 0 to 100. Better indicated by higher values.	57 (1 study) 6 months	⊕⊕⊕⊖ MODERATE1, due to risk of bias		The mean qol: role-physical in the control groups was -8	The mean qol: role-physical in the intervention groups was 30.1 higher (7.86 to 52.34 higher)
QoL: Social functioning SF-36. Scale from: 0 to 100. Better indicated by higher values.	57 (1 study) 6 months	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		The mean qol: social functioning in the control groups was 0.3	The mean qol: social functioning in the intervention groups was 1 higher (13.93 lower to 15.93 higher)
QoL: Role-emotional SF-36. Scale from: 0 to 100. Better indicated by	57 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of		The mean qol: role-emotional in the control groups was	The mean qol: role-emotional in the intervention groups was

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo	Risk difference with T4 (95% CI)
higher values.	6 months	bias, imprecision		2.6	25.1 higher (1.72 to 48.48 higher)
QoL: Mental health SF-36. Scale from: 0 to 100. Better indicated by higher values.	57 (1 study) 6 months	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		The mean qol: mental health in the control groups was 5.6	The mean qol: mental health in the intervention groups was 5.4 lower (18.85 lower to 8.05 higher)
QoL: Vitality SF-36. Scale from: 0 to 100. Better indicated by higher values.	57 (1 study) 6 months	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		The mean qol: vitality in the control groups was 0.2	The mean qol: vitality in the intervention groups was 2.5 lower (14.08 lower to 9.08 higher)
QoL: Bodily pain SF-36. Scale from: 0 to 100. Better indicated by higher values.	57 (1 study) 6 months	⨁⨁⨁⊖ MODERATE1, 2 due to risk of bias		The mean qol: bodily pain in the control groups was -4.6	The mean qol: bodily pain in the intervention groups was 24.3 higher (15.95 to 32.65 higher)
Depression BDI (final values & change scores). Scale from: 0 to 63. Better indicated by lower values.	117 (2 studies) 3-6 months	⨁⨁⨁⊖ MODERATE2 due to risk of bias		The mean depression in the control groups was 4.88	The mean depression in the intervention groups was 0.12 lower (2.6 lower to 2.36 higher)
Hypothyroid symptoms Zulewski, other sign and symptom scale. Multiple scales. Better indicated by lower values.	98 (2 studies) 6-12 months	⨁⨁⨁⨁ HIGH		The mean hypothyroid symptoms in the control groups was 2.05	The mean hypothyroid symptoms in the intervention groups was 0.03 standard deviations lower (-0.43 lower to 0.37 higher)
TSH suppression (<0.4 mIU/L) cases	160 (2 studies) 3 months	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias,	OR 8.12 (2.28 to	0 per 1000	80 more per 1000 (from 30 more to 130 more)3

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo	Risk difference with T4 (95% CI)
		imprecision	28.89)		

1 Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs
2 Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
3 zero events in both arms of one study and one arm of one study

Table 4: Clinical evidence summary: T4 vs Placebo in older adults

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo in older adults	Risk difference with T4 (95% CI)
QoL EQ-5D. Scale from: -0.59 to 1. Better indicated by higher values.	638 (1 study) 12 months	⨁⨁⨁⨀ MODERATE1 due to imprecision		The mean qol in the control groups was 0.85	The mean qol in the intervention groups was 0.02 lower (0.05 lower to 0.01 higher)
QoL EQ VAS. Scale from: 0 to 100. Better indicated by higher values.	638 (1 study) 12 months	⨁⨁⨁⨁ HIGH		The mean qol in the control groups was 77.4	The mean qol in the intervention groups was 0.1 lower (2.38 lower to 2.18 higher)
Hypothyroid symptoms ThyPRO-hypothyroidism. Scale from: 0 to 100.	638 (1 study) 12 months	⨁⨁⨁⨁ HIGH		The mean hypothyroid symptoms in the control groups was 16.7	The mean hypothyroid symptoms in the intervention groups was 0.1 lower (2.77 lower to 2.57 higher)
Hypethyroid symptoms ThyPRO hyperthyroidism. Scale from: 0 to 100.	638 (1 study) 12 months	⨁⨁⨁⨁ HIGH		The mean hyperthyroid symptoms in the control groups was 10.3	The mean hyperthyroid symptoms in the intervention groups was 0.2 higher (1.52 lower to 1.92 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo in older adults	Risk difference with T4 (95% CI)
1 Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs					

See Appendix F: for full GRADE tables.

1.5 Economic evidence

1.5.1 Included studies

No relevant health economic studies were identified.

1.5.2 Excluded studies

No health economic studies that were relevant to this question were excluded due to assessment of limited applicability or methodological limitations.

See also the health economic study selection flow chart in appendix G.

1.5.3 Health economic modelling

This area was not prioritised for new cost-effectiveness analysis.

1.5.4 Resource costs

Relevant unit costs are provided below to aid consideration of cost effectiveness.

Table 4: UK costs of subclinical thyroid treatment

Drug	Daily dose	Cost - month	Cost – annual
Levothyroxine (T_4), 100µg	100µg	£1.34	£16.03
Levothyroxine (T_4) , 75µg	75µg	£3.33	£40.02
Levothyroxine (T_4) , 50µg	50µg	£1.34	£16.03
Levothyroxine (T_4), 25µg	25µg	£2.14	£25.68

Source: BNF, Date, December 2017[17.655]

1.6 Evidence statements

1.6.1 Clinical evidence statements

1.6.1.1 Levothyroxine vs placebo in adults

No clinically important difference was identified for hypothyroid dependent quality of life (1 study, moderate quality), quality of life measures: general health, physical functioning, social functioning (1 study, very low quality), for depression (2 studies, moderate quality), hypothyroid symptoms (2 studies, high quality) and TSH suppression (2 studies, very low quality).

There was a clinically important benefit of levothyroxine for quality of life measures: role physical functioning and bodily pain (1 study, moderate quality), and role emotional functioning (1 study, low quality).

There was a clinically important harm of levothyroxine for quality of life-mental health and vitality (1 study, very low quality).

No evidence was identified for mortality; cardiovascular morbidity-ischemic heart disease; heart failure; arrhythmias; osteoporosis; impaired cognitive function; experience of care; healthcare contacts; growth.

1.6.1.2 Levothyroxine vs placebo in older adults

No clinically important difference was identified for quality of life measured by the EQ-5D (1 study, moderate quality), quality of life measured by the EQ VAS, hypothyroid symptoms and hyperthyroid symptoms (1 study, high quality).

No evidence was identified for mortality; cardiovascular morbidity-ischemic heart disease; heart failure; arrhythmias; osteoporosis; impaired cognitive function; depression; experience of care; healthcare contacts; growth; TSH suppression.

1.6.2 Health economic evidence statements

- No relevant economic evaluations were identified.

1.7 The committee's discussion of the evidence

1.7.1 Interpreting the evidence

1.7.1.1 The outcomes that matter most

The committee agreed that the critical outcomes for this review were mortality and quality of life. Important outcomes were cardiovascular morbidity, arrhythmias, osteoporosis, impaired cognitive function, depression, experience of care, healthcare contacts, symptom scores, growth and TSH suppression.

No clinical evidence was found for mortality. Thus, it was agreed that decision making would be based on quality of life and the important outcomes for which evidence was available.

1.7.1.2 The quality of the evidence

The quality of the evidence ranged from very low to high, being of very low quality for the majority of outcomes. Evidence was typically downgraded for risk of bias which was often attributed to selection bias. Across comparisons, evidence for certain outcomes was also downgraded for imprecision. Overall, the studies included in this review were of relatively short term follow-up, with participants followed up for up to 12 months, during which time they received the interventions.

Levothyroxine vs placebo in adults

The quality of evidence for the use of levothyroxine compared to placebo in adults ranged from very low to high, the majority being of very low quality. The evidence was downgraded mostly due to risk of bias and occasionally due to imprecision. Studies relative to the adult age stratum had a follow up ranging from three to 12 months.

Levothyroxine vs placebo in older adults

The quality of the majority of the evidence for the use of levothyroxine compared to placebo in older adults was high with the exception of one outcome for which the quality of the evidence was moderate and downgraded due to imprecision. Within this comparison, participants of the older age stratum were followed up for 12 months.

1.7.1.3 Benefits and harms

Levothyroxine vs placebo in adults

The evidence showed there was a clinically important benefit with levothyroxine compared to placebo for three quality of life domains: role-physical, role-emotional and bodily pain.

The evidence also showed that there was a clinically important harm with levothyroxine compared to placebo in terms of two quality of life domains, mental health and vitality.

No clinically important difference was found as a result of levothyroxine treatment compared to placebo for hypothyroid dependent quality of life, three quality of life domains: general health, physical functioning and social functioning, depression, hypothyroid symptoms and TSH suppression.

The committee noted that the absence of a clinically important effect could be at least partially attributed to the relatively short-term follow up periods of the studies included in this comparison, which ranged from three to 12 months. They felt that this was likely to be insufficient to observe a clinical difference that is likely to emerge later after treatment. Particularly in regards to depression and symptoms, there was agreement that a longer follow up would be required to draw conclusions about the effect of treatment with levothyroxine for adults with subclinical hypothyroidism.

It was also noted that the dosing strategies of some studies included low doses of levothyroxine that did not reflect current UK practice (100 µg/d), and that this may undermine the effect of levothyroxine on the outcomes measured in the current evidence.

The committee noted the variability in the baseline TSH levels of patients in the studies included in this review. They specified that a TSH greater than 10 would be much less likely to normalise than a lower TSH and agreed on the appropriateness of using this cut off to determine treatment with levothyroxine.

It was raised that an overreliance on TSH levels in decision making about treatment that is most often the case in clinical practice may be problematic, and that other factors, including patients' symptomatology are to influence their need for treatment. The committee felt that a trial period of treatment of 6 months would be appropriate for symptomatic patients with TSH lower than the 10 cut-off.

The importance of making recommendations for both providing but also stopping treatment, in cases where no apparent benefit in symptoms is achieved was emphasised. There was agreement that whether or not TSH returns to normal is a factor indicating the success of treatment but that symptoms are also important.

The committee raised that the presence of antibodies may also influence the likelihood of TSH to return to normal. Within this context, the committee agreed on the importance of considering factors including antibody status and previous thyroid surgery that may suggest an underlying thyroid disease when it comes to the decision of whether or not to offer treatment for subclinical hypothyroidism.

Levothyroxine vs placebo in older adults

Compared to placebo, treatment with levothyroxine did not lead to a clinically important difference in two separate quality of life measures, hypothyroid symptoms and hyperthyroid symptoms.

The committee noted that the evidence for the use of levothyroxine in older adults was underpinned by one study and that a considerable proportion of the participants in this study were asymptomatic and identified based on incidental findings. The committee agreed that it was plausible that the benefits of treatment would be greater in those who had symptoms at baseline.

1.7.2 Cost effectiveness and resource use

There was no health economic evidence identified for this review question, therefore recommendations were based on consensus around treating subclinical hypothyroidism. Unit

costs were presented for different doses of levothyroxine as found in the included clinical studies to aid the committee members in their qualitative judgement in regards to the cost effectiveness.

Levothyroxine vs placebo in adults

Although the evidence about the quality of life was mixed the committee thought by targeting treatment at specific groups that were more likely to benefit, they were maximising the likelihood of treatment being cost effective. Overall, better targeting whom gets treatment compared to what is currently done, people getting treatment according to their TSH levels, is likely to be cost saving. Firstly, because only those who need treatment will be considered, this reduces the number of people being treated unnecessarily, and secondly, people who are considered for treatment will receive a 6 month trial of treatment, after which they will be re-assessed and if no improvement is seen treatment can be stopped. This reduces prescriptions, unnecessary continuation of treatment, compliance issues, and costs.

Furthermore, the committee noted that the current practice for treating subclinical hypothyroidism is done by giving 100µcg daily of levothyroxine tablet (£1.34 per month, BNF, December 2017), which is the cheapest treatment option.

Levothyroxine vs placebo in older adults

The committee did not recommend treatment with levothyroxine for older adults, other than when the TSH was above the reference range but lower than 10 mIU/litre, which is in line with current practice.

1.7.3 Other factors the committee took into account

The committee acknowledged that patients often request treatments with selenium or iodine supplementation. Considering the lack of evidence in regards to those treatments and the frequency with which patients request them, the committee agreed to make research recommendations to investigate their effectiveness for treating subclinical hypothyroidism.

References

1. Abreu IM, Lau E, de Sousa Pinto B, Carvalho D. Subclinical hypothyroidism: to treat or not to treat, that is the question! A systematic review with meta-analysis on lipid profile. Endocrine Connections. 2017; 6(3):188-199

2. Aghili R, Khamseh ME, Malek M, Hadian A, Baradaran HR, Najafi L et al. Changes of subtests of Wechsler Memory Scale and cognitive function in subjects with subclinical hypothyroidism following treatment with levothyroxine. Archives of Medical Science. 2012; 8(6):1096-101

3. Akintola AA, Jansen SW, van Bodegom D, van der Grond J, Westendorp RG, de Craen AJ et al. Subclinical hypothyroidism and cognitive function in people over 60 years: a systematic review and meta-analysis. Frontiers in Aging Neuroscience. 2015; 7:150

4. Angermayr L, Clar C. Iodine supplementation for preventing iodine deficiency disorders in children. Cochrane Database of Systematic Reviews 2004, Issue 2. Art. No.: CD003819. DOI: 10.1002/14651858.CD003819.pub2.

5. Appelhof BC, Peeters RP, Wiersinga WM, Visser TJ, Wekking EM, Huyser J et al. Polymorphisms in type 2 deiodinase are not associated with well-being, neurocognitive functioning, and preference for combined thyroxine/3,5,3'-triiodothyronine therapy. Journal of Clinical Endocrinology and Metabolism. 2005; 90(11):6296-9

6. BMJ Group and the Royal Pharmaceutical Society of Great Britain. British National Formulary. Available from: https://www.evidence.nhs.uk/formulary/bnf/current Last accessed: 04 April 2017

7. Cabral MD, Teixeira P, Soares D, Leite S, Salles E, Waisman M. Effects of thyroxine replacement on endothelial function and carotid artery intima-media thickness in female patients with mild subclinical hypothyroidism. Clinics (Sao Paulo, Brazil). 2011; 66(8):1321-8

8. Caraccio N, Ferrannini E, Monzani F. Lipoprotein profile in subclinical hypothyroidism: response to levothyroxine replacement, a randomized placebo-controlled study. Journal of Clinical Endocrinology and Metabolism. 2002; 87(4):1533-8

9. Caraccio N, Natali A, Sironi A, Baldi S, Frascerra S, Dardano A et al. Muscle metabolism and exercise tolerance in subclinical hypothyroidism: a controlled trial of levothyroxine. Journal of Clinical Endocrinology and Metabolism. 2005; 90(7):4057-62

10. Cassio A, Cacciari E, Cicognani A, Damiani G, Missiroli G, Corbelli E et al. Treatment for congenital hypothyroidism: thyroxine alone or thyroxine plus triiodothyronine? Pediatrics. 2003; 111(5 Pt 1):1055-60

11. Cerbone M, Capalbo D, Wasniewska M, Alfano S, Raso GM, Oliviero U et al. Effects of L-thyroxine treatment on early markers of atherosclerotic disease in children with subclinical hypothyroidism. European Journal of Endocrinology. 2016; 175(1):11-19

12. Cooper DS, Halpern R, Wood LC, Levin AA, Ridgway EC. L-Thyroxine therapy in subclinical hypothyroidism. A double-blind, placebo-controlled trial. Annals of Internal Medicine. 1984; 101(1):18-24

13. Fadeyev VV, Sytch J, Kalashnikov V, Rojtman A, Syrkin A, Melnichenko G. Levothyroxine replacement therapy in patients with subclinical hypothyroidism and coronary artery disease. Endocrine Practice. 2006; 12(1):5-17

14. Feller M, Snel M, Moutzouri E, Bauer DC, de Montmollin M, Aujesky D et al. Association of Thyroid Hormone Therapy With Quality of Life and Thyroid-Related Symptoms in Patients With Subclinical Hypothyroidism: A Systematic Review and Meta-analysis. JAMA. 2018; 320(13):1349-1359

15. Ineck BA, Ng TM. Effects of subclinical hypothyroidism and its treatment on serum lipids. Annals of Pharmacotherapy. 2003; 37(5):725-30

16. Iqbal A, Jorde R, Figenschau Y. Serum lipid levels in relation to serum thyroid-stimulating hormone and the effect of thyroxine treatment on serum lipid levels in subjects with subclinical hypothyroidism: the Tromso Study. Journal of Internal Medicine. 2006; 260(1):53-61

17. Joint Formulary Committee. British National Formulary (BNF) December 2017 update. 2017. Available from: http://www.bnf.org.uk Last accessed: 01/03/2018

18. Jorde R, Waterloo K, Storhaug H, Nyrnes A, Sundsfjord J, Jenssen TG. Neuropsychological function and symptoms in subjects with subclinical hypothyroidism and the effect of thyroxine treatment. Journal of Clinical Endocrinology and Metabolism. 2006; 91(1):145-53

19. Kong WM, Sheikh MH, Lumb PJ, Naoumova RP, Freedman DB, Crook M et al. A 6-month randomized trial of thyroxine treatment in women with mild subclinical hypothyroidism. American Journal of Medicine. 2002; 112(5):348-54

20. Koroglu BK, Bagci O, Ersoy IH, Aksu O, Balkarli A, Alanoglu E et al. Effects of Levothyroxine Treatment on Cardiovascular Risk Profile and Carotid Intima Media Thickness in Patients with Subclinica Hypothyroidism. Acta Endocrinologica. 2012; 8(3):433-442

21. Li X, Meng Z, Jia Q, Ren X. Effect of L-thyroxine treatment versus a placebo on serum lipid levels in patients with sub-clinical hypothyroidism. Biomedical Reports. 2016; 5(4):443-449

22. Mainenti MR, Vigario PS, Teixeira PF, Maia MD, Oliveira FP, Vaisman M. Effect of levothyroxine replacement on exercise performance in subclinical hypothyroidism. Journal of Endocrinological Investigation. 2009; 32(5):470-3

23. Martins RM, Fonseca RH, Duarte MM, Reuters VS, Ferreira MM, Almeida C et al. Impact of subclinical hypothyroidism treatment in systolic and diastolic cardiac function. Arquivos Brasileiros de Endocrinologia e Metabologia. 2011; 55(7):460-7

24. Meier C, Staub JJ, Roth CB, Guglielmetti M, Kunz M, Miserez AR et al. TSH-controlled L-thyroxine therapy reduces cholesterol levels and clinical symptoms in subclinical hypothyroidism: a double blind, placebo-controlled trial (Basel Thyroid Study). Journal of Clinical Endocrinology and Metabolism. 2001; 86(10):4860-6

25. Monzani F, Caraccio N, Kozakowa M, Dardano A, Vittone F, Virdis A et al. Effect of levothyroxine replacement on lipid profile and intima-media thickness in subclinical hypothyroidism: a double-blind, placebo- controlled study. Journal of Clinical Endocrinology and Metabolism. 2004; 89(5):2099-106

26. Monzani F, Di Bello V, Caraccio N, Bertini A, Giorgi D, Giusti C et al. Effect of levothyroxine on cardiac function and structure in subclinical hypothyroidism: a double blind, placebo-controlled study. Journal of Clinical Endocrinology and Metabolism. 2001; 86(3):1110-5

27. Nagasaki T, Inaba M, Yamada S, Shirakawa K, Nagata Y, Kumeda Y et al. Decrease of brachial-ankle pulse wave velocity in female subclinical hypothyroid patients during

normalization of thyroid function: a double-blind, placebo-controlled study. European Journal of Endocrinology. 2009; 160(3):409-15

28. Najafi L, Malek M, Hadian A, Ebrahim Valojerdi A, Khamseh ME, Aghili R. Depressive symptoms in patients with subclinical hypothyroidism--the effect of treatment with levothyroxine: a double-blind randomized clinical trial. Endocrine Research. 2015; 40(3):121-6

29. National Institute for Health and Care Excellence. Developing NICE guidelines: the manual [updated October 2018]. London. National Institute for Health and Care Excellence, 2014. Available from: http://www.nice.org.uk/article/PMG20/chapter/1%20Introduction%20and%20overview

30. Ng SM, Anand D, Weindling AM. High versus low dose of initial thyroid hormone replacement for congenital hypothyroidism. Cochrane Database of Systematic Reviews 2009, Issue 1. Art. No.: CD006972. DOI: 10.1002/14651858.CD006972.pub2.

31. Nystrom E, Caidahl K, Fager G, Wikkelso C, Lundberg PA, Lindstedt G. A double-blind cross-over 12-month study of L-thyroxine treatment of women with 'subclinical' hypothyroidism. Clinical Endocrinology. 1988; 29(1):63-75

32. Parle J, Roberts L, Wilson S, Pattison H, Roalfe A, Haque MS et al. A randomized controlled trial of the effect of thyroxine replacement on cognitive function in community-living elderly subjects with subclinical hypothyroidism: the Birmingham Elderly Thyroid study. Journal of Clinical Endocrinology and Metabolism. 2010; 95(8):3623-32

33. Pollock MA, Sturrock A, Marshall K, Davidson KM, Kelly CJ, McMahon AD et al. Thyroxine treatment in patients with symptoms of hypothyroidism but thyroid function tests within the reference range: randomised double blind placebo controlled crossover trial. BMJ. 2001; 323(7318):891-5

34. Razvi S, Ingoe L, Keeka G, Oates C, McMillan C, Weaver JU. The beneficial effect of L-thyroxine on cardiovascular risk factors, endothelial function, and quality of life in subclinical hypothyroidism: randomized, crossover trial. Journal of Clinical Endocrinology and Metabolism. 2007; 92(5):1715-23

35. Reuters VS, Almeida Cde P, Teixeira Pde F, Vigario Pdos S, Ferreira MM, Castro CL et al. Effects of subclinical hypothyroidism treatment on psychiatric symptoms, muscular complaints, and quality of life. Arquivos Brasileiros de Endocrinologia e Metabologia. 2012; 56(2):128-36

36. Ross DS. Bone density is not reduced during the short-term administration of levothyroxine to postmenopausal women with subclinical hypothyroidism: a randomized, prospective study. American Journal of Medicine. 1993; 95(4):385-8

37. Ruggeri RM, Trimarchi F, Biondi B. L-Thyroxine replacement therapy in the frail elderly: A challenge in clinical practice. European Journal of Endocrinology. 2017; 177(4):R199-R217

38. Segna D, Bauer DC, Feller M, Schneider C, Fink HA, Aubert CE et al. Association between subclinical thyroid dysfunction and change in bone mineral density in prospective cohorts. Journal of Internal Medicine. 2017; 283(1):56-72

39. Shatynska-Mytsyk I, Rodrigo L, Cioccocioppo R, Petrovic D, Lakusic N, Compostella L et al. The impact of thyroid hormone replacement therapy on left ventricular diastolic function in patients with subclinical hypothyroidism. Journal of Endocrinological Investigation. 2016; 39(6):709-713

40. Stott DJ, Rodondi N, Kearney PM, Ford I, Westendorp RGJ, Mooijaart SP et al. Thyroid hormone therapy for older adults with subclinical hypothyroidism. New England Journal of Medicine. 2017; 376(26):2534-2544

41. Sun J, Yao L, Fang Y, Yang R, Chen Y, Yang K et al. Relationship between Subclinical Thyroid Dysfunction and the Risk of Cardiovascular Outcomes: A Systematic Review and Meta-Analysis of Prospective Cohort Studies. International Journal of Endocrinology Print. 2017; 2017:8130796

42. Taylor PN, Okosieme OE, Dayan CM, Lazarus JH. Impact of iodine supplementation in mild-to-moderate iodine deficiency: Systematic review and meta-analysis. European Journal of Endocrinology. 2014; 170(1):R1-R15

43. Teixeira PF, Reuters VS, Ferreira MM, Almeida CP, Reis FA, Melo BA et al. Treatment of subclinical hypothyroidism reduces atherogenic lipid levels in a placebo-controlled double-blind clinical trial. Hormone and Metabolic Research. 2008; 40(1):50-5

44. Valizadeh M, Seyyedmajidi MR, Momtazi S, Musavi Nasab N. The efficacy of combined levothyroxine plus liothyronine with levothyroxine alone in primary hypothyroidism: a randomized controlled trial. Iranian journal of endocrinology and metabolism. 2009; 10(5):465-471

45. Villar HCCE, Saconato H, Valente O, Atallah ÁN. Thyroid hormone replacement for subclinical hypothyroidism. Cochrane Database of Systematic Reviews 2007, Issue 3. Art. No.: CD003419. DOI: 10.1002/14651858.CD003419.pub2.

46. Wasniewska M, Corrias A, Aversa T, Valenzise M, Mussa A, De Martino L et al. Comparative evaluation of therapy with L-thyroxine versus no treatment in children with idiopathic and mild subclinical hypothyroidism. Hormone research in pdiatrics. 2012; 77(6):376-81

47. Yazici M, Gorgulu S, Sertbas Y, Erbilen E, Albayrak S, Yildiz O et al. Effects of thyroxin therapy on cardiac function in patients with subclinical hypothyroidism: index of myocardial performance in the evaluation of left ventricular function. International Journal of Cardiology. 2004; 95(2-3):135-43

48. Yetmis M, Kazancioglu R, Erkoc R, Tukek T, Peru C, Cikrkcioglu MA. Changes in lipid profile and body mass index in patients with subclinical hypothyroidism: Evaluation of L-Thyroxine treatment. Haseki Tip Bulteni. 2011; 49(4):131-135

49. Zhao M, Liu L, Wang F, Yuan Z, Zhang X, Xu C et al. A Worthy Finding: Decrease in Total Cholesterol and Low-Density Lipoprotein Cholesterol in Treated Mild Subclinical Hypothyroidism. Thyroid. 2016; 26(8):1019-29

50. Zhao T, Chen B, Zhou Y, Wang X, Zhang Y, Wang H et al. Effect of levothyroxine on the progression of carotid intima-media thickness in subclinical hypothyroidism patients: a meta-analysis. BMJ Open. 2017; 7(10):e016053

51. Zhu L, Bai X, Teng WP, Shan ZY, Wang WW, Fan CL et al. [Effects of selenium supplementation on antibodies of autoimmune thyroiditis]. Chinese Medical Journal (Taipei). 2012; 92(32):2256-2260

Appendices

Appendix A: Review protocols

Table 5:

ID	Field	Content
I	Review question	What is the clinical and cost effectiveness of treating subclinical hypothyroidism?
II	Type of review question	Intervention A review of health economic evidence related to the same review question was conducted in parallel with this review. For details see the health economic review protocol for this NICE guideline.
III	Objective of the review	Determine the most clinically and cost effective way to treat subclinical hypothyroidism
IV	Eligibility criteria – population / disease / condition / issue / domain	People diagnosed with subclinical hypothyroidism (TSH greater than upper limit of context specific normal range, T3/T4 within normal range)
V	Eligibility criteria – intervention(s) / exposure(s) / prognostic factor(s)	T3 T4 – initiation at high dose T4 – initiation via gradual titration Combination of T3 & T4 Natural thyroid extract (mammalian only) Iodine supplementation Selenium supplementation Placebo
VI	Eligibility criteria – comparator(s) / control or reference (gold) standard	Any of above vs any other, in isolation or combination
VII	Outcomes and prioritisation	**Critical** • Mortality (dichotomous, ≥1 year) • Quality of life (continuous) **Important** • Cardiovascular morbidity - ischaemic heart disease, heart failure (dichotomous) • Arhythmias (dichotomous) • Osteoporosis (dichotomous) • Impaired cognitive function/neurodevelopment for children (dichotomous) • Depression (dichotomous) • Patient/family/carer experience of care (continuous) • Healthcare contacts (rates/dichotomous) • Symptom scores (continuous) • Growth (continuous) • TSH suppression (dichotomous)

		Minimum duration as for the minimum duration for inclusion of studies unless specified.
VIII	Eligibility criteria – study design	• RCTs only • Blinded comparisons prioritised, non-blinded comparisons only considered if blinded unavailable on an intervention by intervention basis • Minimum treatment duration of 3 months • Crossover studies included
IX	Other inclusion exclusion criteria	• Europe based studies only for selenium supplementation to maintain representative selenium status in trial populations to UK population • Studies in areas/populations of severe iodine deficiency excluded for iodine supplementation excluded • Studies in pregnant women excluded
X	Proposed sensitivity / subgroup analysis, or meta-regression	**Stratifications** • Age – children (under 4), children (4-18), adults (>18-65), older adults (>65) • TSH at initiation of treatment – TSH 2.5-<5 U/ml, 5-<10 U/ml, 10 or more U/ml (only applicable to treatment naïve) **Subgroup analyses** • Age subdivisions (18-50, 50-65, 65-80, >85) • T4 treatment strategy (liquid vs pill, daily vs weekly) • Children on dietary restrictions vs general diet
XI	Selection process – duplicate screening / selection / analysis	• No duplicate screening was deemed necessary for this question, for more information please see the separate Methods report for this guideline.
XII	Data management (software)	• Pairwise meta-analyses were performed using Cochrane Review Manager (RevMan5). • GRADEpro was used to assess the quality of evidence for each outcome. • Endnote was used for bibliography, citations, sifting and reference management
XIII	Information sources – databases and dates	• Medline, Embase and the Cochrane Library
XIV	Identify if an update	Not an update
XV	Author contacts	https://www.nice.org.uk/guidance/indevelopment/gid-ng10074
XVI	Highlight if amendment to previous protocol	Not an amendment
XVII	Search strategy – for one database	For details please see Appendix B:.
XVIII	Data collection process – forms /	A standardised evidence table format was used, and published as an appendix of the evidence report.

	duplicate	
XIX	Data items – define all variables to be collected	For details please see evidence tables in Appendix D: (clinical evidence tables) or Appendix H: (health economic evidence tables).
XX	Methods for assessing bias at outcome / study level	Standard study checklists were used to critically appraise individual studies. For details please see section 6.2 of Developing NICE guidelines: the manual The risk of bias across all available evidence was evaluated for each outcome using an adaptation of the 'Grading of Recommendations Assessment, Development and Evaluation (GRADE) toolbox' developed by the international GRADE working group http://www.gradeworkinggroup.org/
XXI	Criteria for quantitative synthesis	For details please see section 6.4 of Developing NICE guidelines: the manual.
XXII	Methods for quantitative analysis – combining studies and exploring (in)consistency	For details please see the separate Methods report for this guideline.
XXIII	Meta-bias assessment – publication bias, selective reporting bias	For details please see section 6.2 of Developing NICE guidelines: the manual.
XXIV	Confidence in cumulative evidence	For details please see sections 6.4 and 9.1 of Developing NICE guidelines: the manual.
XXV	Rationale / context – what is known	For details please see the introduction to the evidence review.
XXVI	Describe contributions of authors and guarantor	A multidisciplinary committee [to add link to history page of the guideline after publication] developed the evidence review. The committee was convened by the National Guideline Centre (NGC) and chaired by Sarah Fishburn in line with section 3 of Developing NICE guidelines: the manual. Staff from NGC undertook systematic literature searches, appraised the evidence, conducted meta-analysis and cost-effectiveness analysis where appropriate, and drafted the evidence review in collaboration with the committee. For details please see Developing NICE guidelines: the manual.
XXVII	Sources of funding / support	NGC is funded by NICE and hosted by the Royal College of Physicians.
XXVIII	Name of sponsor	NGC is funded by NICE and hosted by the Royal College of Physicians.
XXIX	Roles of sponsor	NICE funds NGC to develop guidelines for those working in the NHS, public health and social care in England.
XXX	PROSPERO registration number	Not registered

1

1 **Table 6: Health economic review protocol**

Review question	All questions – health economic evidence
Objectives	To identify health economic studies relevant to any of the review questions.
Search criteria	• Populations, interventions and comparators must be as specified in the clinical review protocol above. • Studies must be of a relevant health economic study design (cost–utility analysis, cost-effectiveness analysis, cost–benefit analysis, cost–consequences analysis, comparative cost analysis). • Studies must not be a letter, editorial or commentary, or a review of health economic evaluations. (Recent reviews will be ordered although not reviewed. The bibliographies will be checked for relevant studies, which will then be ordered.) • Unpublished reports will not be considered unless submitted as part of a call for evidence. • Studies must be in English.
Search strategy	A health economic study search will be undertaken using population-specific terms and a health economic study filter – see Appendix B: below.
Review strategy	Studies not meeting any of the search criteria above will be excluded. Studies published before 2003, abstract-only studies and studies from non-OECD countries or the USA will also be excluded. Each remaining study will be assessed for applicability and methodological limitations using the NICE economic evaluation checklist which can be found in appendix H of Developing NICE guidelines: the manual (2014).[29] **Inclusion and exclusion criteria** • If a study is rated as both 'Directly applicable' and with 'Minor limitations' then it will be included in the guideline. A health economic evidence table will be completed and it will be included in the health economic evidence profile. • If a study is rated as either 'Not applicable' or with 'Very serious limitations' then it will usually be excluded from the guideline. If it is excluded then a health economic evidence table will not be completed and it will not be included in the health economic evidence profile. • If a study is rated as 'Partially applicable', with 'Potentially serious limitations' or both then there is discretion over whether it should be included. **Where there is discretion** The health economist will make a decision based on the relative applicability and quality of the available evidence for that question, in discussion with the guideline committee if required. The ultimate aim is to include health economic studies that are helpful for decision-making in the context of the guideline and the current NHS setting. If several studies are considered of sufficiently high applicability and methodological quality that they could all be included, then the health economist, in discussion with the committee if required, may decide to include only the most applicable studies and to selectively exclude the remaining studies. All studies excluded on the basis of applicability or methodological limitations will be listed with explanation in the excluded health economic studies appendix below. The health economist will be guided by the following hierarchies. *Setting:* • UK NHS (most applicable). • OECD countries with predominantly public health insurance systems (for example, France, Germany, Sweden). • OECD countries with predominantly private health insurance systems (for example, Switzerland).

	• Studies set in non-OECD countries or in the USA will be excluded before being assessed for applicability and methodological limitations. *Health economic study type:* • Cost–utility analysis (most applicable). • Other type of full economic evaluation (cost–benefit analysis, cost-effectiveness analysis, cost–consequences analysis). • Comparative cost analysis. • Non-comparative cost analyses including cost-of-illness studies will be excluded before being assessed for applicability and methodological limitations. *Year of analysis:* • The more recent the study, the more applicable it will be. • Studies published in 2003 or later but that depend on unit costs and resource data entirely or predominantly from before 2003 will be rated as 'Not applicable'. • Studies published before 2003 will be excluded before being assessed for applicability and methodological limitations. *Quality and relevance of effectiveness data used in the health economic analysis:* • The more closely the clinical effectiveness data used in the health economic analysis match with the outcomes of the studies included in the clinical review the more useful the analysis will be for decision-making in the guideline.

1
2

Appendix B: Literature search strategies

The literature searches for this review are detailed below and complied with the methodology outlined in Developing NICE guidelines: the manual 2014, updated 2018 https://www.nice.org.uk/guidance/pmg20/resources/developing-nice-guidelines-the-manual-pdf-72286708700869

For more detailed information, please see the Methodology Review. [Add cross reference after publication]

B.1 Clinical search literature search strategy

Searches were constructed using a PICO framework where population (P) terms were combined with Intervention (I) and in some cases Comparison (C) terms. Outcomes (O) are rarely used in search strategies for interventions as these concepts may not be well described in title, abstract or indexes and therefore difficult to retrieve. Search filters were applied to the search where appropriate.

Database	Dates searched	Search filter used
Medline (OVID)	1946 – 07 January 2019	Exclusions Randomised controlled trials Systematic review studies
Embase (OVID)	1974 – 07 January 2019	Exclusions Randomised controlled trials Systematic review studies
The Cochrane Library (Wiley)	Cochrane Reviews to 2019 Issue 1 or 12 CENTRAL to 2019 Issue 1 or 12 DARE, and NHSEED to 2015 Issue 2 of 4 HTA to 2016 Issue 2 of 4	None

Medline (Ovid) search terms

1.	exp thyroid diseases/
2.	hyperthyroid*.ti,ab.
3.	hypothyroid*.ti,ab.
4.	thyrotoxicosis.ti,ab.
5.	(thyroid adj3 (swell* or dysfunction* or enlarg* or nodule* or node* or disease* or condition* or disorder*)).ti,ab.
6.	or/1-5
7.	letter/
8.	editorial/
9.	news/
10.	exp historical article/
11.	Anecdotes as Topic/
12.	comment/
13.	case report/
14.	(letter or comment*).ti.
15.	or/7-14

16.	randomized controlled trial/ or random*.ti,ab.
17.	15 not 16
18.	animals/ not humans/
19.	exp Animals, Laboratory/
20.	exp Animal Experimentation/
21.	exp Models, Animal/
22.	exp Rodentia/
23.	(rat or rats or mouse or mice).ti.
24.	or/17-23
25.	6 not 24
26.	limit 25 to English language
27.	((iodine or selenium) adj2 supplement*).ti,ab.
28.	(desiccated adj3 (thyroid or hormone* or extract or extracts)).ti,ab.
29.	(thyroid adj2 (extract or extracts)).ti,ab.
30.	(natural adj4 thyroid).ti,ab.
31.	(natural adj3 (extract or extracts)).ti,ab.
32.	armour*.ti,ab.
33.	(thyroxine or levothyroxine or liothyronine or triiodothyronine or tri-iodothyronine).ti,ab.
34.	Thyroxine/ or Triiodothyronine/
35.	(T3 or T4).ti,ab.
36.	(TSH or thyroid stimulating hormone or thyrotropin).ti,ab.
37.	or/27-36
38.	26 and 37
39.	randomized controlled trial.pt.
40.	controlled clinical trial.pt.
41.	randomi#ed.ti,ab.
42.	placebo.ab.
43.	randomly.ti,ab.
44.	Clinical Trials as topic.sh.
45.	trial.ti.
46.	or/39-45
47.	Meta-Analysis/
48.	exp Meta-Analysis as Topic/
49.	(meta analy* or metanaly* or metaanaly* or meta regression).ti,ab.
50.	((systematic* or evidence*) adj3 (review* or overview*)).ti,ab.
51.	(reference list* or bibliograph* or hand search* or manual search* or relevant journals).ab.
52.	(search strategy or search criteria or systematic search or study selection or data extraction).ab.
53.	(search* adj4 literature).ab.
54.	(medline or pubmed or cochrane or embase or psychlit or psyclit or psychinfo or psycinfo or cinahl or science citation index or bids or cancerlit).ab.
55.	cochrane.jw.
56.	((multiple treatment* or indirect or mixed) adj2 comparison*).ti,ab.
57.	or/47-56
58.	38 and (46 or 57)

1 **Embase (Ovid) search terms**

1.	exp thyroid disease/
2.	hyperthyroid*.ti,ab.
3.	hypothyroid*.ti,ab.
4.	thyrotoxicosis.ti,ab.
5.	(thyroid adj3 (swell* or dysfunction* or enlarg* or nodule* or node* or disease* or condition* or disorder*)).ti,ab.
6.	or/1-5
7.	letter.pt. or letter/
8.	note.pt.
9.	editorial.pt.
10.	case report/ or case study/
11.	(letter or comment*).ti.
12.	or/7-11
13.	randomized controlled trial/ or random*.ti,ab.
14.	12 not 13
15.	animal/ not human/
16.	nonhuman/
17.	exp Animal Experiment/
18.	exp Experimental Animal/
19.	animal model/
20.	exp Rodent/
21.	(rat or rats or mouse or mice).ti.
22.	or/14-21
23.	6 not 22
24.	limit 23 to English language
25.	((iodine or selenium) adj2 supplement*).ti,ab.
26.	(desiccated adj3 (thyroid or hormone* or extract or extracts)).ti,ab.
27.	(thyroid adj2 (extract or extracts)).ti,ab.
28.	(natural adj4 thyroid).ti,ab.
29.	armour*.ti,ab.
30.	*thyroxine/ or *levothyroxine/ or *liothyronine/ or *triiodothyronine/
31.	(thyroxine or levothyroxine or liothyronine or triiodothyronine or tri-iodothyronine).ti,ab.
32.	(T3 or T4).ti,ab.
33.	(TSH or thyroid stimulating hormone or thyrotropin).ti,ab.
34.	*thyrotropin/
35.	or/25-34
36.	24 and 35
37.	random*.ti,ab.
38.	factorial*.ti,ab.
39.	(crossover* or cross over*).ti,ab.
40.	((doubl* or singl*) adj blind*).ti,ab.
41.	(assign* or allocat* or volunteer* or placebo*).ti,ab.
42.	crossover procedure/
43.	single blind procedure/

44.	randomized controlled trial/
45.	double blind procedure/
46.	or/37-45
47.	systematic review/
48.	meta-analysis/
49.	(meta analy* or metanaly* or metaanaly* or meta regression).ti,ab.
50.	((systematic* or evidence*) adj3 (review* or overview*)).ti,ab.
51.	(reference list* or bibliograph* or hand search* or manual search* or relevant journals).ab.
52.	(search strategy or search criteria or systematic search or study selection or data extraction).ab.
53.	(search* adj4 literature).ab.
54.	(medline or pubmed or cochrane or embase or psychlit or psyclit or psychinfo or psycinfo or cinahl or science citation index or bids or cancerlit).ab.
55.	cochrane.jw.
56.	((multiple treatment* or indirect or mixed) adj2 comparison*).ti,ab.
57.	or/47-56
58.	36 and (46 or 57)

Cochrane Library (Wiley) search terms

#1.	MeSH descriptor: [Thyroid Diseases] explode all trees
#2.	hyperthyroid*:ti,ab
#3.	hypothyroid*:ti,ab
#4.	thyrotoxicosis:ti,ab
#5.	(thyroid near/3 (swell* or dysfunction* or enlarg* or nodule* or node* or disease* or condition* or disorder*)):ti,ab
#6.	(or #1-#5)
#7.	((iodine or selenium) near/2 supplement*):ti,ab
#8.	(desiccated near/3 (thyroid or hormone* or extract or extracts)):ti,ab
#9.	(thyroid near/2 (extract or extracts)):ti,ab
#10.	(natural near/4 thyroid):ti,ab
#11.	(natural near/3 (extract or extracts)):ti,ab
#12.	armour*:ti,ab
#13.	(thyroxine or levothyroxine or liothyronine or triiodothyronine or tri-iodothyronine):ti,ab
#14.	MeSH descriptor: [Thyroxine] explode all trees
#15.	MeSH descriptor: [Triiodothyronine] explode all trees
#16.	(T3 or T4):ti,ab
#17.	(TSH or thyroid stimulating hormone or thyrotropin):ti,ab
#18.	(or #7-#17)
#19.	#6 and #18

B.2 Health Economics literature search strategy

Health economic evidence was identified by conducting a broad search relating to a thyroid disease population in NHS Economic Evaluation Database (NHS EED – this ceased to be updated after March 2015) and the Health Technology Assessment database (HTA) with no date restrictions. NHS EED and HTA databases are hosted by the Centre for Research and

Dissemination (CRD). Additional searches were run on Medline and Embase for health economics, economic modelling and quality of life studies.

Table 7: Database date parameters and filters used

Database	Dates searched	Search filter used
Medline	2014 – 07 January 2019	Exclusions Health economics studies Health economics modelling studies Quality of life studies
Embase	2014 – 07 January 2019	Exclusions Health economics studies Health economics modelling studies Quality of life studies
Centre for Research and Dissemination (CRD)	HTA - Inception – 07 January 2019 NHSEED - Inception to March 2015	None

Medline (Ovid) search terms

1.	exp thyroid diseases/
2.	hyperthyroid*.ti,ab.
3.	hypothyroid*.ti,ab.
4.	thyrotoxicosis.ti,ab.
5.	(thyroid adj3 (swell* or dysfunction* or enlarg* or nodule* or node* or disease* or condition* or disorder*)).ti,ab.
6.	or/1-5
7.	letter/
8.	editorial/
9.	news/
10.	exp historical article/
11.	Anecdotes as Topic/
12.	comment/
13.	case report/
14.	(letter or comment*).ti.
15.	or/7-14
16.	randomized controlled trial/ or random*.ti,ab.
17.	15 not 16
18.	animals/ not humans/
19.	exp Animals, Laboratory/
20.	exp Animal Experimentation/
21.	exp Models, Animal/
22.	exp Rodentia/
23.	(rat or rats or mouse or mice).ti.
24.	or/17-23
25.	6 not 24
26.	limit 25 to English language

27.	Economics/
28.	Value of life/
29.	exp "Costs and Cost Analysis"/
30.	exp Economics, Hospital/
31.	exp Economics, Medical/
32.	Economics, Nursing/
33.	Economics, Pharmaceutical/
34.	exp "Fees and Charges"/
35.	exp Budgets/
36.	budget*.ti,ab.
37.	cost*.ti.
38.	(economic* or pharmaco?economic*).ti.
39.	(price* or pricing*).ti,ab.
40.	(cost* adj2 (effective* or utilit* or benefit* or minimi* or unit* or estimat* or variable*)).ab.
41.	(financ* or fee or fees).ti,ab.
42.	(value adj2 (money or monetary)).ti,ab.
43.	or/27-42
44.	exp models, economic/
45.	*Models, Theoretical/
46.	*Models, Organizational/
47.	markov chains/
48.	monte carlo method/
49.	exp Decision Theory/
50.	(markov* or monte carlo).ti,ab.
51.	econom* model*.ti,ab.
52.	(decision* adj2 (tree* or analy* or model*)).ti,ab.
53.	or/44-52
54.	quality-adjusted life years/
55.	sickness impact profile/
56.	(quality adj2 (wellbeing or well being)).ti,ab.
57.	sickness impact profile.ti,ab.
58.	disability adjusted life.ti,ab.
59.	(qal* or qtime* or qwb* or daly*).ti,ab.
60.	(euroqol* or eq5d* or eq 5*).ti,ab.
61.	(qol* or hql* or hqol* or h qol* or hrqol* or hr qol*).ti,ab.
62.	(health utility* or utility score* or disutilit* or utility value*).ti,ab.
63.	(hui or hui1 or hui2 or hui3).ti,ab.
64.	(health* year* equivalent* or hye or hyes).ti,ab.
65.	discrete choice*.ti,ab.
66.	rosser.ti,ab.
67.	(willingness to pay or time tradeoff or time trade off or tto or standard gamble*).ti,ab.
68.	(sf36* or sf 36* or short form 36* or shortform 36* or shortform36*).ti,ab.
69.	(sf20 or sf 20 or short form 20 or shortform 20 or shortform20).ti,ab.
70.	(sf12* or sf 12* or short form 12* or shortform 12* or shortform12*).ti,ab.

71.	(sf8* or sf 8* or short form 8* or shortform 8* or shortform8*).ti,ab.
72.	(sf6* or sf 6* or short form 6* or shortform 6* or shortform6*).ti,ab.
73.	or/54-72
74.	26 and (43 or 53 or 73)

Embase (Ovid) search terms

1.	exp thyroid diseases/
2.	hyperthyroid*.ti,ab.
3.	hypothyroid*.ti,ab.
4.	thyrotoxicosis*.ti,ab.
5.	(thyroid adj3 (swell* or dysfunction* or enlarg* or nodule* or node* or disease* or condition* or disorder*)).ti,ab.
6.	or/1-5
7.	letter.pt. or letter/
8.	note.pt.
9.	editorial.pt.
10.	case report/ or case study/
11.	(letter or comment*).ti.
12.	or/7-11
13.	randomized controlled trial/ or random*.ti,ab.
14.	12 not 13
15.	animal/ not human/
16.	nonhuman/
17.	exp Animal Experiment/
18.	exp Experimental Animal/
19.	animal model/
20.	exp Rodent/
21.	(rat or rats or mouse or mice).ti.
22.	or/14-21
23.	6 not 22
24.	limit 23 to English language
25.	health economics/
26.	exp economic evaluation/
27.	exp health care cost/
28.	exp fee/
29.	budget/
30.	funding/
31.	budget*.ti,ab.
32.	cost*.ti.
33.	(economic* or pharmaco?economic*).ti.
34.	(price* or pricing*).ti,ab.
35.	(cost* adj2 (effective* or utilit* or benefit* or minimi* or unit* or estimat* or

	variable*)).ab.
36.	(financ* or fee or fees).ti,ab.
37.	(value adj2 (money or monetary)).ti,ab.
38.	or/25-37
39.	statistical model/
40.	exp economic aspect/
41.	39 and 40
42.	*theoretical model/
43.	*nonbiological model/
44.	stochastic model/
45.	decision theory/
46.	decision tree/
47.	monte carlo method/
48.	(markov* or monte carlo).ti,ab.
49.	econom* model*.ti,ab.
50.	(decision* adj2 (tree* or analy* or model*)).ti,ab.
51.	or/41-50
52.	quality adjusted life year/
53.	"quality of life index"/
54.	short form 12/ or short form 20/ or short form 36/ or short form 8/
55.	sickness impact profile/
56.	(quality adj2 (wellbeing or well being)).ti,ab.
57.	sickness impact profile.ti,ab.
58.	disability adjusted life.ti,ab.
59.	(qal* or qtime* or qwb* or daly*).ti,ab.
60.	(euroqol* or eq5d* or eq 5*).ti,ab.
61.	(qol* or hql* or hqol* or h qol* or hrqol* or hr qol*).ti,ab.
62.	(health utility* or utility score* or disutilit* or utility value*).ti,ab.
63.	(hui or hui1 or hui2 or hui3).ti,ab.
64.	(health* year* equivalent* or hye or hyes).ti,ab.
65.	discrete choice*.ti,ab.
66.	rosser.ti,ab.
67.	(willingness to pay or time tradeoff or time trade off or tto or standard gamble*).ti,ab.
68.	(sf36* or sf 36* or short form 36* or shortform 36* or shortform36*).ti,ab.
69.	(sf20 or sf 20 or short form 20 or shortform 20 or shortform20).ti,ab.
70.	(sf12* or sf 12* or short form 12* or shortform 12* or shortform12*).ti,ab.
71.	(sf8* or sf 8* or short form 8* or shortform 8* or shortform8*).ti,ab.
72.	(sf6* or sf 6* or short form 6* or shortform 6* or shortform6*).ti,ab.
73.	or/52-72
74.	24 and (38 or 51 or 73)

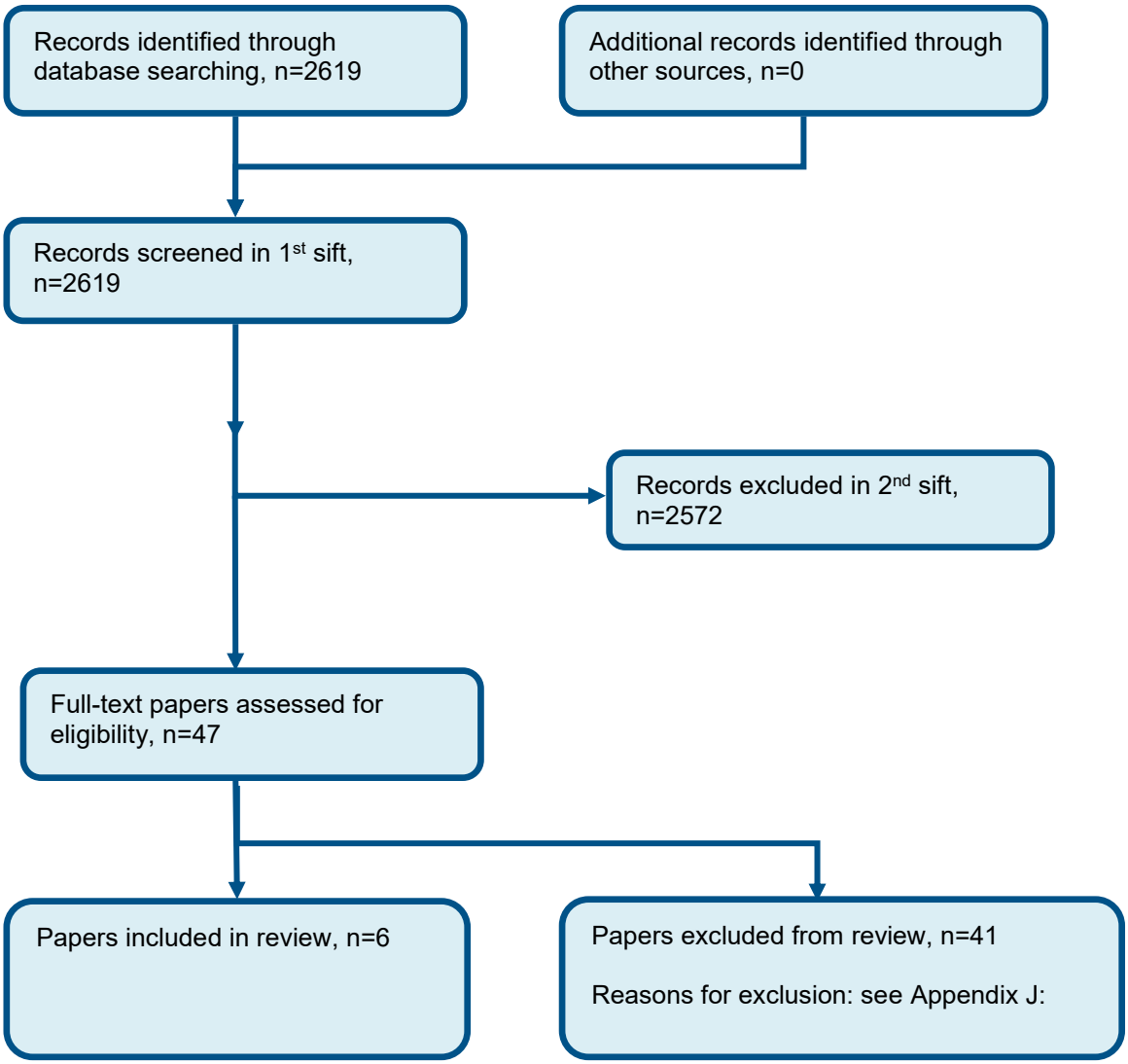
1 **NHS EED and HTA (CRD) search terms**

#1.	MeSH DESCRIPTOR Thyroid Diseases EXPLODE ALL TREES
#2.	hyperthyroid*
#3.	hypothyroid*
#4.	thyrotoxicosis*
#5.	(thyroid adj3 (swell* or dysfunction* or enlarg* or nodule* or node* or disease* or condition* or disorder*))
#6.	#1 OR #2 OR #3 OR #4 or #5

2

Appendix C: Clinical evidence selection

Figure 1: Flow chart of clinical study selection for the review of management of subclinical hypothyroidism



Appendix D: Clinical evidence tables

Study	Kong 2002[19]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=40)
Countries and setting	Conducted in United Kingdom; Setting: Not stated
Line of therapy	1st line
Duration of study	Intervention + follow up: 6 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Thyroid function test
Stratum	Naive - TSH 5-10
Subgroup analysis within study	Not applicable
Inclusion criteria	Normal serum free thyroxin (FT4) 0.8-16 ng/dL and TSH levels between 5-10 µU/mL
Exclusion criteria	History of previous thyroid disease, psychiatric disorder or anticipated pregnancy.
Recruitment/selection of patients	Consecutive
Age, gender and ethnicity	Age - Mean (SD): T4: 53 (3); Placebo: 45(4). Gender (M:F): 0/40. Ethnicity: Not stated
Further population details	
Indirectness of population	No indirectness
Interventions	(n=23) Intervention 1: T4 only - T4 - high dose start. 50-100 µg daily (depending on TSH >5 µU/mL) . Duration 6 months. Concurrent medication/care: Not stated. Indirectness: No indirectness Further details: 1. T4 dosing: 2. T4 formulations: (n=17) Intervention 2: Placebo. ascorbic acid, 5 mg daily. Duration 6 months. Concurrent medication/care: Not stated. Indirectness: No indirectness Further details: 1. T4 dosing: 2. T4 formulations:
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: T4 - HIGH DOSE START versus PLACEBO Protocol outcome 1: Symptom scores - Actual outcome for Naive - TSH 5-10: Symptom scores at 6 months; Group 1: mean 2.6 (SD 1.4); n=20, Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: -- ; Group 1 Number missing: 3, Reason: unwell on T4, increased fatigue, pregnancy; Group 2 Number missing: 2, Reason: increase fatigue, noncompliance	
Protocol outcomes not reported by the study	Quality of life ; Mortality ; Ischaemic heart disease ; Heart failure ; Arrhythmia ; Osteoporosis ; Impaired cognitive function ; Depression ; Experience of care ; Healthcare contacts ; Growth ; TSH suppression

Study	Meier 2001[24]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=66)
Countries and setting	Conducted in Switzerland; Setting: University Hospital Basel
Line of therapy	1st line
Duration of study	Intervention + follow up: 48 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: examination, full medical and endocrine work-up
Stratum	Overall
Subgroup analysis within study	Post-hoc subgroup analysis: by TSH concentration (>12 mlU/liter; n=13)
Inclusion criteria	Age 18-75 years, TSH level more than 5mlU/L on 2 consecutive blood tests, exaggerated TSH response of more than 35 mlU/L after oral TRH stimulation, free T4 within normal range, good general health as assessed by a full medical and endocrine work-up
Exclusion criteria	coronary heart disease, pituitary hypothalamic disorders or other non-thyroidal illnesses ; thyroid hormone medication up to 3 months before enrollment; lipid lowering agents within 6 months before enrollment; obvious or suspected poor compliance
Recruitment/selection of patients	Not specified
Age, gender and ethnicity	Age - Mean (SD): T4: 57.1 (10.34), Placebo: 57.1 (10.91). Gender (M:F): 0/63. Ethnicity: Not specified
Further population details	
Indirectness of population	No indirectness
Interventions	(n=33) Intervention 1: T4 only - T4 - titrated dose start. 25, 50, 75, 199 or 125 µg daily, adapted every 6 weeks in the first 24 weeks to achieve euthyroid TSH (0.1-4.0 mlU/liter), mean dose at end of study : 85.5, SD 4.3. Duration 48 weeks. Concurrent medication/care: -. Indirectness: No indirectness Further details: 1. T4 dosing: Not applicable 2. T4 formulations: (n=33) Intervention 2: Placebo. controlled every 6 weeks to ascertain an optimal replacement regimen. Duration 48 weeks. Concurrent medication/care: -. Indirectness: No indirectness Further details: 1. T4 dosing: 2. T4 formulations:
Funding	Academic or government funding (Swiss Research Foundation)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: T4 - TITRATED DOSE START versus PLACEBO

Protocol outcome 1: Symptom scores
- Actual outcome: Clinical signs and symptoms at 48 weeks; Group 1: mean 1.5 (SD 1.11); n=31, Group 2: mean 1.6 (SD 1.13); n=32; Zulewski 0-12 Top=High is poor outcome
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2, Reason: serious medical comorbidities; Group 2 Number missing: 1, Reason: malignant astrocytoma

Protocol outcome 2: TSH suppression
- Actual outcome: TSH suppression below reference range (0.1-4.0 mIU/liter) at 48 weeks; Group 1: 0/31, Group 2: 0/32
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2, Reason: serious medical comorbidities; Group 2 Number missing: 1, Reason: malignant astrocytoma

Protocol outcomes not reported by the study	Quality of life ; Mortality ; Ischaemic heart disease ; Heart failure ; Arrhythmia ; Osteoporosis ; Impaired cognitive function ; Depression ; Experience of care ; Healthcare contacts ; Growth

Study	Najafi 2015[28]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=60)
Countries and setting	Conducted in Iran; Setting: Endocrine Research Centre, Institute of Endocrinology and Metabolism
Line of therapy	1st line
Duration of study	Intervention + follow up: 3 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: detailed history and physical examination by expert physician
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	subclinical hypothyroidism, serum TSH > 4.5 mIU/L, normal free T4 (0.8-2 ng/dl), positive anti-TPO-Ab
Exclusion criteria	History of endocrine or autoimmune disease other than subclinical hypothyroidism, history of thyroid hormone, or corticosteroids replacement in the previous 2 months. Diabetes, mellitus, heart failure, chronic liver or pulmonary disorder, history of head trauma, seizure, known psychological or mental disorders, pregnancy.
Recruitment/selection of patients	Referral to the Thyroid Clinic of the Institute of Endocrinology and Metabolism.
Age, gender and ethnicity	Age - Mean (SD): 34 (10). Gender (M:F): 9/51. Ethnicity: Not stated
Further population details	
Indirectness of population	No indirectness
Interventions	(n=30) Intervention 1: T4 only - T4 - high dose start. 100 µg. Duration 12 weeks. Concurrent medication/care: Not specified. Indirectness: No indirectness Further details: 1. T4 dosing: 2. T4 formulations: (n=30) Intervention 2: Placebo. Placebo tablets. Duration 12 weeks. Concurrent medication/care: Not specified. Indirectness: No indirectness Further details: 1. T4 dosing: 2. T4 formulations:
Funding	Academic or government funding (Tehran University of Medical Sciences (TUMS))
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: T4 - HIGH DOSE START versus PLACEBO	

Protocol outcome 1: Depression - Actual outcome: Depression at 12 weeks; Group 1: mean 12.37 (SD 10.01); n=30, Group 2: mean 11.86 (SD 10.71); n=30; Bech Depression Inventory (BDI) 0-63 Top=High is poor outcome Risk of bias: All domain - Low, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: Serious indirectness, Comments: Continuous analysis; Baseline details: Baseline differences; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 2: TSH suppression - Actual outcome: TSH suppression at 12 weeks; Group 1: 0/30, Group 2: 0/30 Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Quality of life ; Mortality ; Ischaemic heart disease ; Heart failure ; Arrhythmia ; Osteoporosis ; Impaired cognitive function ; Experience of care ; Healthcare contacts ; Symptom scores ; Growth

Study	Razvi 2007[34]
Study type	RCT (Patient randomised; Crossover: No washout)
Number of studies (number of participants)	1 (n=100)
Countries and setting	Conducted in United Kingdom; Setting:
Line of therapy	1st line
Duration of study	Intervention time: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: two thyroid function tests
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Stable SCH (TSH >4 mIU/liter and FT4 levels in the normal reference range), aged 18-80 years, having had at least two thyroid function tests measured at least 3 months apart
Exclusion criteria	previous thyroid disease and its treatment, medications that could cause thyroid hormone dysfunction, diabetes mellitus, serum creatinine greater than 1.36 mg/dl, vascular disease, psychiatric conditions or their treatment and current or previous pregnancy in the last 2 years.
Recruitment/selection of patients	from laboratory database
Age, gender and ethnicity	Age - Mean (SD): 53.8 (12.6). Gender (M:F): 18/ 82. Ethnicity: UK
Further population details	
Indirectness of population	No indirectness
Interventions	(n=100) Intervention 1: T4 only - T4 - high dose start. 100 µg daily. Duration 12 weeks. Concurrent medication/care: -. Indirectness: No indirectness Further details: 1. T4 dosing: 2. T4 formulations: (n=100) Intervention 2: Placebo. 100 µg daily. Duration 12 weeks. Concurrent medication/care: -. Indirectness: No indirectness Further details: 1. T4 dosing: 2. T4 formulations:
Funding	Academic or government funding (National Health Service Research and Development)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: T4 ONLY versus PLACEBO	

Protocol outcome 1: Quality of life - Actual outcome: Hypothyroid dependent QoL at 12 months; Group 1: mean -1.1 (SD 1); n=100, Group 2: mean -1.2 (SD 0.9); n=100; T-QoL (hypothyroid dependent QoL) -3 to 1 Top=High is good outcome Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: dropped out due to side effects Protocol outcome 2: TSH suppression - Actual outcome: TSH suppression below reference (<0.4 mIU/L) at 12 months; Group 1: 10/100, Group 2: 0/100 Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: dropped out due to side effects	
Protocol outcomes not reported by the study	Mortality ; Ischaemic heart disease ; Heart failure ; Arrhythmia ; Osteoporosis ; Impaired cognitive function ; Depression ; Experience of care ; Healthcare contacts ; Symptom scores ; Growth

Study	Reuters 2012[35]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=71)
Countries and setting	Conducted in Brazil; Setting: outpatients
Line of therapy	Mixed line
Duration of study	Intervention + follow up: 6 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: physical examination
Stratum	Overall: Stratified by TSH levels (level 1: >4.0-8.0 µUI/mL, level 2: >8.0-12.0 µUI/mL, level 3: > 12.0 µUI/mL
Subgroup analysis within study	Not applicable
Inclusion criteria	TSH> 4.0 µUI/mL and normal FT4 (0.9-1.8 ng/dL) confirmed by two laboratory serum determinations with minimum interval of six weeks, a confirmation of laboratory euthyroidism for at least a year before development of sHP if sHT developed after hyperthyroidism treatment,
Exclusion criteria	Chronic disease, using drugs that may influence thyroid function, patients with severe psychiatric disturbances, patients who attended school for less than 3 years, using drugs or showing diseases that could influence neuromuscular function
Recruitment/selection of patients	not stated
Age, gender and ethnicity	Age - Mean (SD): T4: 49.3 (10.3); Placebo: 50.7 (11.4). Gender (M:F): 7/62. Ethnicity: Not stated
Further population details	
Indirectness of population	No indirectness
Interventions	(n=35) Intervention 1: T4 only - T4 - high dose start. 25 mcg, 50 mcg or 75 mcg depending on stratification by TSH levels (adjusted at two months). Duration 6 months. Concurrent medication/care: Not reported. Indirectness: No indirectness Further details: 1. T4 dosing: 2. T4 formulations: Comments: Patients progressing to overt hypothyroidism or requiring more than 75 mcg to reach euthyroidism were excluded (n=36) Intervention 2: Placebo. adjusted by TSH levels. Duration 6 months. Concurrent medication/care: Not reported. Indirectness: No indirectness Further details: 1. T4 dosing: 2. T4 formulations:

Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: T4 - HIGH DOSE START versus PLACEBO

Protocol outcome 1: Quality of life
- Actual outcome: Quality of life: General health at 6 months; Group 1: mean 6.5 (SD 13.2); n=25, Group 2: mean 7.4 (SD 18.3); n=32; SF-36 -general health 0-100 Top=High is good outcome
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness; Baseline details: Baseline general health scores not reported; Group 1 Number missing: 10, Reason: disease progression, hyperthyroidism, need to use drug or disease, no adherence; Group 2 Number missing: 4, Reason: disease progression, need to used drug or disease, no adherence
- Actual outcome: Quality of life: Physical functioning at 6 months; Group 1: mean 3.7 (SD 17.2); n=25, Group 2: mean 1.9 (SD 21.6); n=32; SF-35: Physical functionig 0-100 Top=High is good outcome
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness; Baseline details: Baseline general health scores not reported; Group 1 Number missing: 10, Reason: disease progression, hyperthyroidism, need to use drug or disease, no adherence; Group 2 Number missing: 4, Reason: disease progression, need to used drug or disease, no adherence
- Actual outcome: Quality of life: Role-physical at 6 months; Group 1: mean 22.1 (SD 47.5); n=25, Group 2: mean -8 (SD 35.1); n=32; SF-36: Role-physical 0-100 Top=High is good outcome
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness; Baseline details: Baseline general health scores not reported; Group 1 Number missing: 10, Reason: disease progression, hyperthyroidism, need to use drug or disease, no adherence; Group 2 Number missing: 4, Reason: disease progression, need to used drug or disease, no adherence
- Actual outcome: Quality of life: Social functioning at 6 months; Group 1: mean 1.3 (SD 24.4); n=25, Group 2: mean 0.3 (SD 33.1); n=32; SF-36: Social functioning 0-100 Top=High is good outcome
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness; Baseline details: Baseline general health scores not reported; Group 1 Number missing: 10, Reason: disease progression, hyperthyroidism, need to use drug or disease, no adherence; Group 2 Number missing: 4, Reason: disease progression, need to used drug or disease, no adherence
- Actual outcome: Quality of life: Role emotional at 6 months; Group 1: mean 27.7 (SD 47.5); n=25, Group 2: mean 2.6 (SD 40.8); n=32; SF-36: Role-emotional 0-100 Top=High is good outcome
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness; Baseline details: Baseline general health scores not reported; Group 1 Number missing: 10, Reason: disease progression, hyperthyroidism, need to use drug or disease, no adherence; Group 2 Number missing: 4, Reason: disease progression, need to used drug

or disease, no adherence - Actual outcome: Quality of life: Mental health at 6 months; Group 1: mean 0.2 (SD 28.2); n=25, Group 2: mean 5.6 (SD 22.1); n=32 Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness; Baseline details: Baseline general health scores not reported; Group 1 Number missing: 10, Reason: disease progression, hyperthyroidism, need to use drug or disease, no adherence; Group 2 Number missing: 4, Reason: disease progression, need to used drug or disease, no adherence - Actual outcome: Quality of life: Vitality at 6 months; Group 1: mean -2.3 (SD 22.9); n=25, Group 2: mean 0.2 (SD 21.1); n=32; SF-36-Vitality 0-100 Top=High is good outcome Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness; Baseline details: Baseline general health scores not reported; Group 1 Number missing: 10, Reason: disease progression, hyperthyroidism, need to use drug or disease, no adherence; Group 2 Number missing: 4, Reason: disease progression, need to used drug or disease, no adherence - Actual outcome: Quality of life: Bodily pain at 6 months; Group 1: mean 19.7 (SD 15.2); n=25, Group 2: mean -4.6 (SD 16.9); n=32; Sf-36- Bodily pain 0-100 Top=High is good outcome Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness; Baseline details: Baseline general health scores not reported; Group 1 Number missing: 10, Reason: disease progression, hyperthyroidism, need to use drug or disease, no adherence; Group 2 Number missing: 4, Reason: disease progression, need to used drug or disease, no adherence Protocol outcome 2: Depression - Actual outcome: Depressive symptoms at 6 months; Group 1: mean -2.4 (SD 5.8); n=25, Group 2: mean -2.1 (SD 4.8); n=32; Beck Depression Inventory (BDI) 0-63 Top=High is poor outcome Risk of bias: All domain - High, Selection - Very high, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: Serious indirectness, Comments: Continuous analysis; Baseline details: Baseline scores not reported; Group 1 Number missing: 10, Reason: disease progression, hyperthyroidism, need to use drug or disease, no adherence; Group 2 Number missing: 4, Reason: disease progression, need to used drug or disease, no adherence	
Protocol outcomes not reported by the study	Mortality ; Ischaemic heart disease ; Heart failure ; Arrhythmia ; Osteoporosis ; Impaired cognitive function ; Experience of care ; Healthcare contacts ; Symptom scores ; Growth ; TSH suppression

Study	Stott 2017[40]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=737)
Countries and setting	Conducted in Netherlands, Switzerland, United Kingdom; Setting: multiple centres
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 months
Method of assessment of guideline condition	Method of assessment /diagnosis not stated
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Age 65 years or more, persistent subclinical hypothyroidism defined as TSH 4.60-19.99 mIU/L, measured on at least two occasions that were 3 months to 3 years apart, F-T4 within reference range,
Exclusion criteria	current prescription of levothyroxine, antithyroid drugs, amiodarone, or lithium; thyroid surgery or receipt of radioactive iodine within the previous 12 months; dementia; hospitalization for a major illness; elective surgery within previous 4 weeks; an acute coronary heart syndrome within the previous 4 weeks); and terminal illness.
Recruitment/selection of patients	identified from clinical laboratory and general practice databases
Age, gender and ethnicity	Age - Mean (SD): 74.4 (6.3). Gender (M:F): 341/396. Ethnicity: 98% white
Further population details	
Indirectness of population	No indirectness
Interventions	(n=368) Intervention 1: T4 only - T4 - high dose start. 50 µg daily (or 25 if body weight <50 kg or known coronary heart disease), adjusted for goal TSH (0.40-4.59 mIU/L). Duration 12 months. Concurrent medication/care: not specified. Indirectness: No indirectness Further details: 1. T4 dosing: 2. T4 formulations: (n=369) Intervention 2: Placebo. mock adjustment. Duration 12 months. Concurrent medication/care: not specified. Indirectness: No indirectness Further details: 1. T4 dosing: 2. T4 formulations:
Funding	Academic or government funding (European Union FP7)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: T4 - HIGH DOSE START versus PLACEBO Protocol outcome 1: Quality of life - Actual outcome for Treatment naive/general population: health-related quality of life (EQ VAS) at 12 months; Group 1: mean 77.3 (SD 15.6); n=318, Group 2: mean 77.4 (SD 13.7); n=320 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - High, Comments - ; Indirectness of outcome: No indirectness ; Group 1 Number missing: 17, Reason: Not specified; Group 2 Number missing: 14, Reason: Not specified - Actual outcome for Treatment naive/general population: health-related quality of life (EQ-5D) at 12 months; Group 1: mean 0.833 (SD 0.212); n=318, Group 2: mean 0.853 (SD 0.191); n=320; EQ-5D 0.59-1 Top=High is good outcome Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - High, Comments - ; Indirectness of outcome: No indirectness ; Group 1 Number missing: 17, Reason: Not specified; Group 2 Number missing: 14, Reason: Not specified Protocol outcome 2: Symptom scores - Actual outcome for Treatment naive/general population: Hypothyroid symptom scores at 12 months; Group 1: mean 16.6 (SD 16.9); n=318, Group 2: mean 16.7 (SD 17.5); n=320; Thyroid related quality of life patient reported outcome measure (ThyPRO) hypothyroid symptoms score 0-100 Top=High is poor outcome Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - High; Indirectness of outcome: No indirectness ; Baseline details: Not substantial difference in scores at baseline; Group 1 Number missing: 17, Reason: Not specified; Group 2 Number missing: 14, Reason: Not specified - Actual outcome for Treatment naive/general population: Hyperthyroid symptoms scores at 12 months; Group 1: mean 10.5 (SD 10.8); n=318, Group 2: mean 10.3 (SD 11.3); n=320; ThyPRO Hyperthyroid Symptoms 0-100 Top=High is poor outcome; Comments: minimum clinically impotant difference estimated as 9 points Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - High; Indirectness of outcome: No indirectness ; Baseline details: Not substantial difference in scores at baseline; Group 1 Number missing: 17, Reason: Not specified; Group 2 Number missing: 14, Reason: Not specified	
Protocol outcomes not reported by the study	Mortality ; Ischaemic heart disease ; Heart failure ; Arrhythmia ; Osteoporosis ; Impaired cognitive function ; Depression ; Experience of care ; Healthcare contacts ; Growth ; TSH suppression

Appendix E: Forest plots

E.1 Subclinical hypothyroidism – T4 vs placebo in Adults

Figure 3: Quality of life: hypothyroid dependent (T-QoL, -3 to 1, high = good, at 12 months)

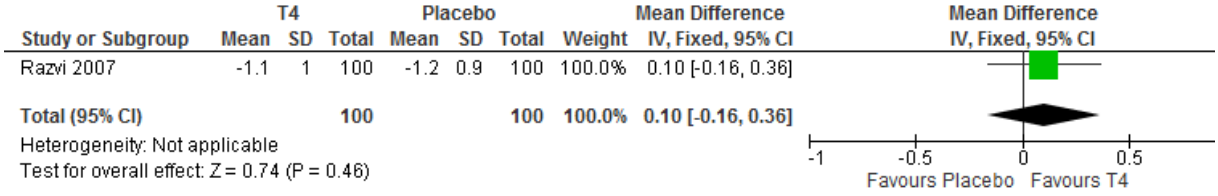


Figure 4: Quality of life: general health (SF-36, 0-100, high=good, at 6 months)

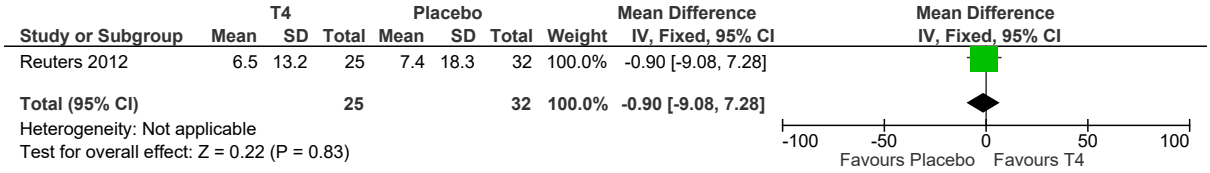


Figure 5: Quality of life: physical functioning (SF-36, 0-100, high=good, at 6 months)

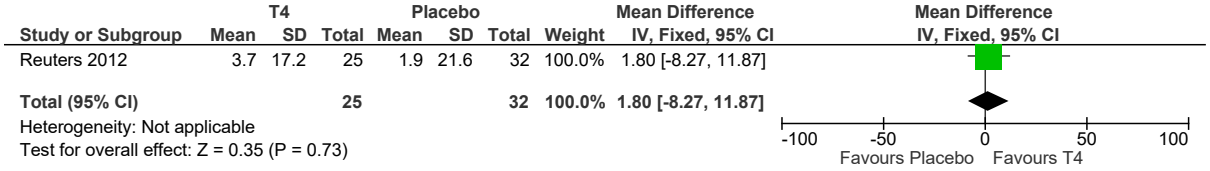


Figure 6: Quality of life: role-physical (SF-36, 0-100, high=good, at 6 months)

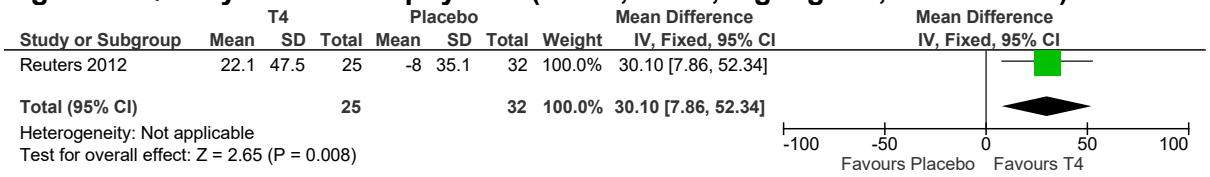


Figure 7: Quality of life: social functioning (SF-36, 0-100, high=good, at 6 months)

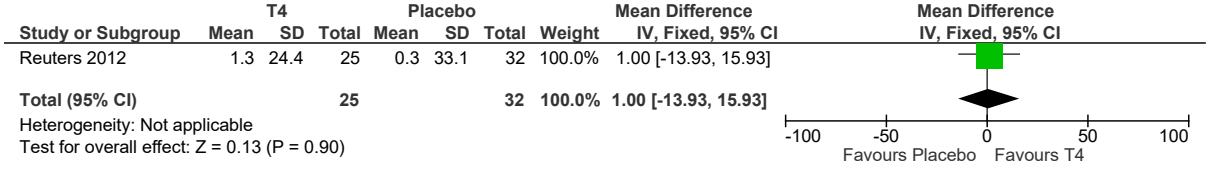
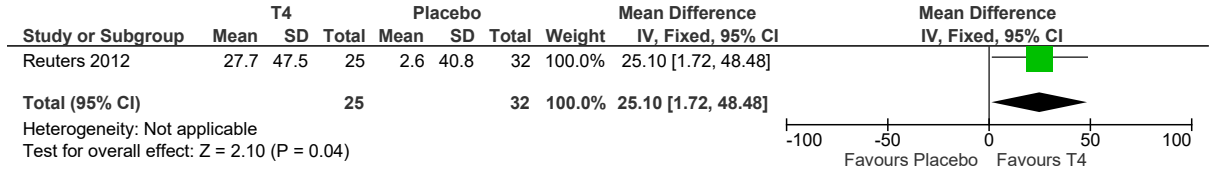
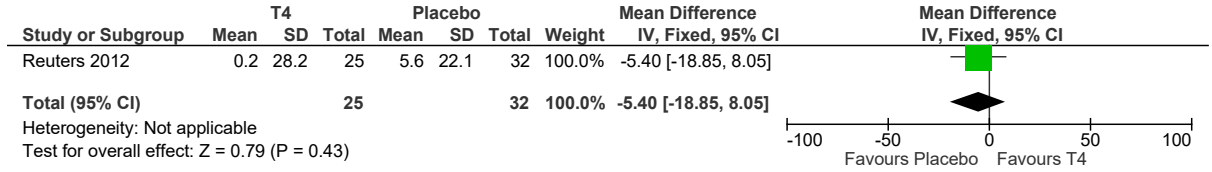


Figure 8: Quality of life: role-emotional (SF-36, 0-100, high=good, at 6 months)



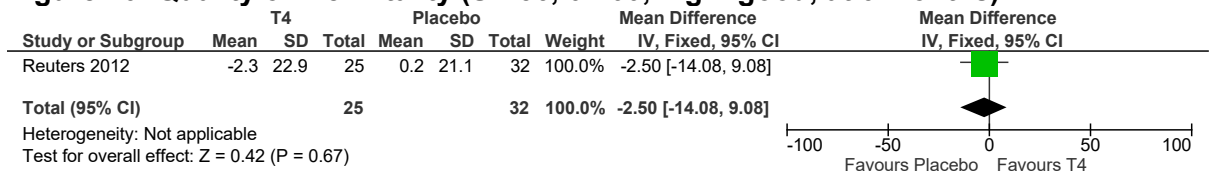
1

Figure 9: Quality of life: mental health (SF-36, 0-100, high=good, at 6 months)



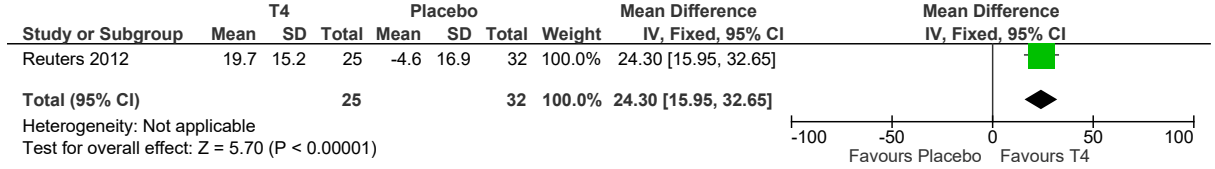
2

Figure 10: Quality of life: vitality (SF-36, 0-100, high=good, at 6 months)



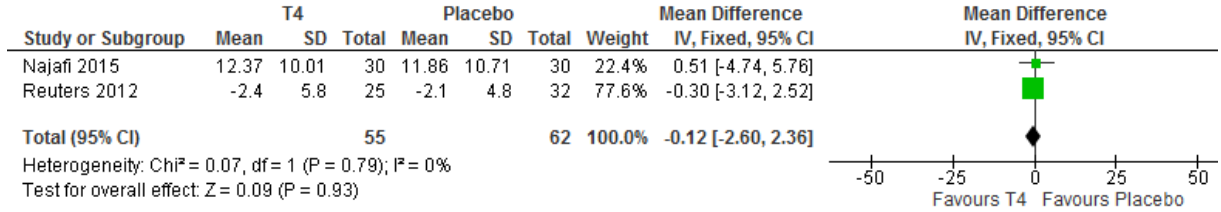
3

Figure 11: Quality of life: bodily pain (SF-36, 0-100, high=good, at 6 months)



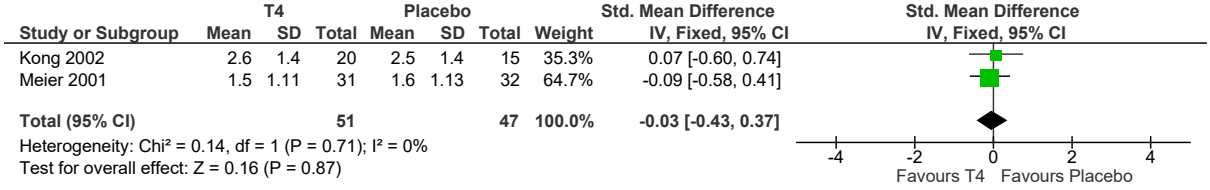
4

Figure 12: Depression (BDI, 0-63, high=poor, at 3-6 months)



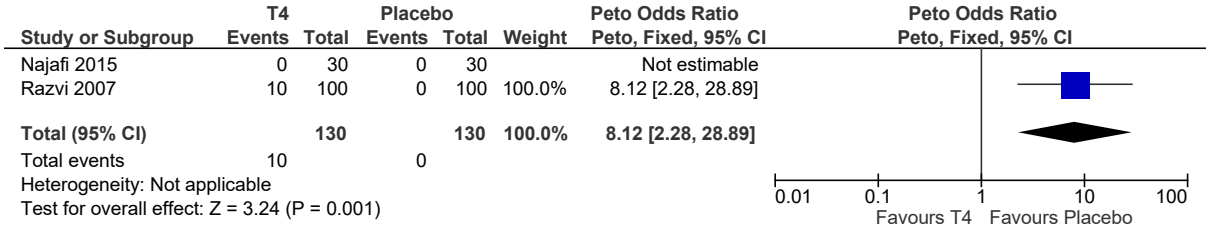
5

Figure 13: Hypothyroid symptoms (Multiple scales, high=poor, at 6-12 months)



1

Figure 14: TSH suppression (> reference, at 3 months)



2

3

4 E.2 Subclinical hypothyroidism - T4 vs placebo for older adults

Figure 15: Quality of life (EQ-5D, -0.59-1, high=good, at 12 months)

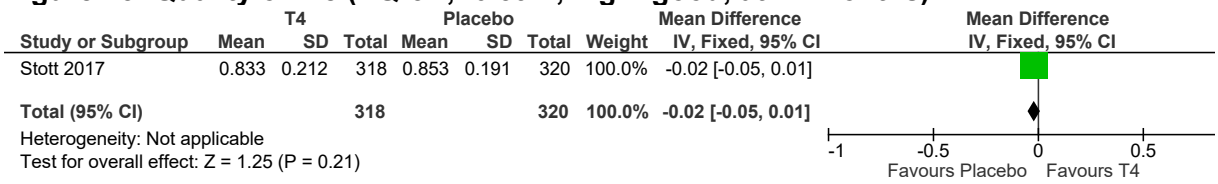
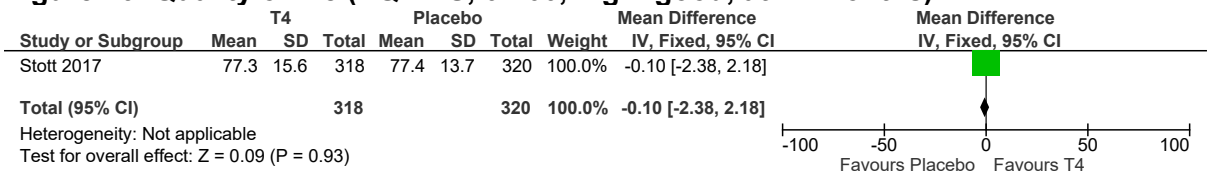
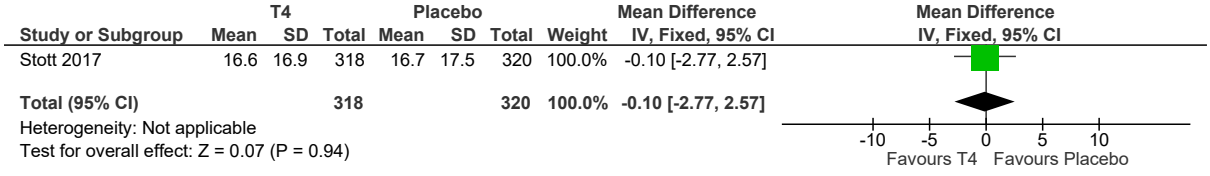


Figure 16: Quality of life (EQ VAS, 0-100, high=good, at 12 months)



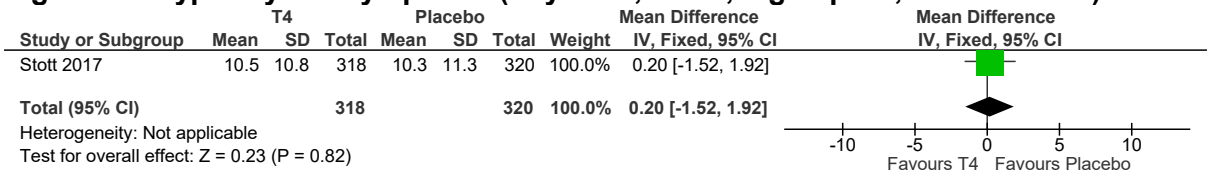
5

Figure 17: Hypothyroid symptoms (ThyPRO, 0-100, high=poor, at 12 months)



6

Figure 18: Hyperthyroid symptoms (Thy-PRO, 0-100, high= poor, at 12 months)



Appendix F: GRADE tables

Table 8: Clinical evidence profile: T4 versus placebo in adults

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	T4	Placebo	Relative (95% CI)	Absolute		
QoL: Hypothyroid-dependent (follow-up 12 months; measured with: T-QoL; range of scores: -3-1; Better indicated by higher values)												
1	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	no serious imprecision	none	100	100	-	MD 0.1 higher (0.16 lower to 0.36 higher)	⊕⊕⊕O MODERATE	CRITICAL
QoL: General health (follow-up 6 months; measured with: SF-36; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	very serious[1]	none	25	32	-	MD 0.9 lower (9.08 lower to 7.28 higher)	⊕OOO VERY LOW	CRITICAL
QoL: Physical functioning (follow-up 6 months; measured with: SF-36; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	very serious[1]	none	25	32	-	MD 1.8 higher (8.27 lower to 11.87 higher)	⊕OOO VERY LOW	CRITICAL
QoL: Role-physical (follow-up 6 months; measured with: SF-36; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	no serious imprecision[1]	none	25	32	-	MD 30.1 higher (7.86 to 52.34 higher)	⊕⊕⊕O MODERATE	CRITICAL
QoL: Social functioning (follow-up 6 months; measured with: SF-36; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	very serious[1]	none	25	32	-	MD 1 higher (13.93 lower to 15.93 higher)	⊕OOO VERY LOW	CRITICAL
QoL: Role-emotional (follow-up 6 months; measured with: SF-36; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	serious[1]	none	25	32	-	MD 25.1 higher (1.72 to 48.48 higher)	⊕⊕OO	CRITICAL

											LOW	
QoL: Mental health (follow-up 6 months; measured with: SF-36; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	very serious[1]	none	25	32	-	MD 5.4 lower (18.85 lower to 8.05 higher)	⊕OOO VERY LOW	CRITICAL
QoL: Vitality (follow-up 3-6 months; measured with: SF-36; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	very serious[1]	none	25	32	-	MD 2.5 lower (14.08 lower to 9.08 higher)	⊕OOO VERY LOW	CRITICAL
QoL: Bodily pain (follow-up 6 months; measured with: SF-36; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	no serious imprecision[1]	none	25	32	-	MD 24.3 higher (15.95 to 32.65 higher)	⊕⊕⊕O MODERATE	CRITICAL
Depression (follow-up 3-6 months; measured with: BDI (final values & change scores); range of scores: 0-63; Better indicated by lower values)												
2	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	no serious imprecision	none	55	62	-	MD 0.12 lower (2.6 lower to 2.36 higher)	⊕⊕⊕O MODERATE	IMPORTANT
Hypothyroid symptoms (follow-up 6-12 months; measured with: Zulewski,signs and symptoms of hypothyroidism; Multiple scales; Better indicated by lower values)												
2	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	51	47	-	SMD 0.03 lower (0.43 lower to 0.37 higher)	⊕⊕⊕⊕ HIGH	IMPORTANT
TSH suppression (<0.4 mIU/L) (follow-up 3 months; assessed with: cases)												
2	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	very serious[1]	none	10/130 (7.7%)	0%	OR 8.12 (2.28 to 28.89)	80 more per 1000 (from 0.03 more to 0.13 more)[3]	⊕OOO VERY LOW	IMPORTANT

[1] Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
[2] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs
[3] zero events in both arms of one study and one arm of one study

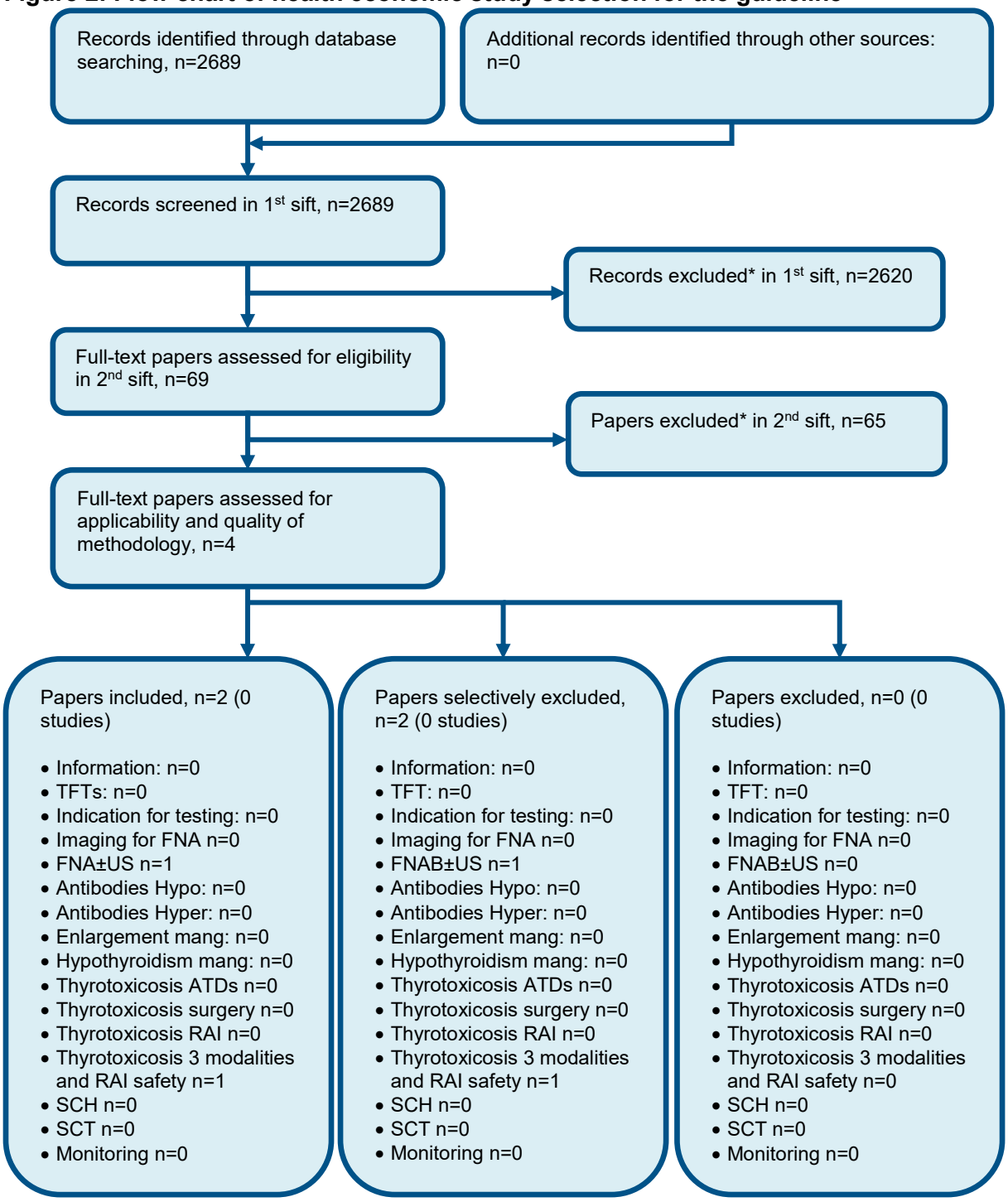
Table 9: Clinical evidence profile: T4 versus placebo in older adults

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	T4	Placebo in older adults	Relative (95% CI)	Absolute		
QoL (follow-up 12 months; measured with: EQ-5D; range of scores: -0.59-1; Better indicated by higher values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	318	320	-	MD 0.02 lower (0.05 lower to 0.01 higher)	⊕⊕⊕O MODERATE	CRITICAL
QoL (follow-up 12 months; measured with: EQ VAS; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	318	320	-	MD 0.1 lower (2.38 lower to 2.18 higher)	⊕⊕⊕⊕ HIGH	CRITICAL
Hypothyroid symptoms (follow-up 12 months; measured with: ThyPRO-hypothyroidism; range of scores: 0-100; Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	318	320	-	MD 0.1 lower (2.77 lower to 2.57 higher)	⊕⊕⊕⊕ HIGH	IMPORTANT
Hypethyroid symptoms (follow-up 12 months; measured with: ThyPRO hyperthyroidism; range of scores: 0-100; Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	318	320	-	MD 0.2 higher (1.52 lower to 1.92 higher)	⊕⊕⊕⊕ HIGH	IMPORTANT

[1] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

Appendix G: Health economic evidence selection

Figure 2: Flow chart of health economic study selection for the guideline



* Non-relevant population, intervention, comparison, design or setting; non-English language
TFT; thyroid function test, FNA; fine-needle aspiration, US; ultrasound, RAI; radioactive iodine, ATDs; antithyroid drugs, Mang; management, SCH; Subclinical hypothyroidism, SCT; Subclinical thyrotoxicosis.

1

2

Appendix H: Health economic evidence tables

None

Appendix I: Health economic analysis

None

Appendix J: Excluded studies

J.1 Excluded clinical studies

Table 10: Studies excluded from the clinical review

Study	Exclusion reason
Abreu 2017[1]	References checked
Aghili 2012[2]	No usable outcomes
Akintola 2015[3]	Systematic review-references checked
Angermayr 2004[4]	Not review population
Appelhof 2005[5]	Not review population
Cabral 2011[7]	No usable outcomes
Caraccio 2002[8]	No usable outcomes
Caraccio 2005[9]	No usable outcomes
Cassio 2003[10]	Less than minimum duration
Cerbone 2016[11]	Inappropriate comparison. Incorrect interventions. No usable outcomes. inappropriate control group
Cooper 1984[12]	No usable outcomes matching protocol
Fadeyev 2006[13]	no usable outcomes matching protocol
Feller 2018[14]	SR, references checked
Ineck 2003[15]	SR, references checked
Iqbal 2006[16]	No usable outcomes
Jorde 2006[18]	No usable outcomes
Koroglu 2012[20]	No usable outcomes
Li 2016[21]	References checked
Mainenti 2009[22]	Incorrect interventions. Inappropriate comparison. No placebo
Martins 2011[23]	No usable outcomes matching protocol
Monzani 2001[26]	No usable outcomes
Monzani 2004[25]	No usable outcome matching protocol
Nagasaki 2009[27]	No usable outcomes
Ng 2009[30]	Not review population. References checked
Nystrom 1988[31]	No usable outcomes matching protocol
Parle 2010[32]	No usable outcomes
Pollock 2001[33]	Not review population
Ross 1993[36]	No usable outcomes matching protocol. No control group
Ruggeri 2017[37]	References checked
Segna 2017[38]	Systematic review: study designs inappropriate. Inappropriate comparison. Incorrect interventions
Shatynska-mytsyk 2016[39]	Non-randomised controlled study
Sun 2017[41]	Systematic review: study designs inappropriate. References checked
Taylor 2014[42]	Supplementation study in iodine deficient country. References checked
Teixeira 2008[43]	no usable outcomes matching protocol
Valizadeh 2009[44]	Not review population
Wasniewska 2012[46]	Incorrect interventions. no usable outcomes matching protocol
Yazici 2004[47]	No usable outcomes

Yetmis 2011[48]	Not in English
Zhao 2016[49]	No usable outcomes
Zhao 2017[50]	References checked
Zhu 2012[51]	Not in English

J.2 Excluded health economic studies

None

Appendix K: Research recommendations

K.1

Research question: What is the clinical and cost effectiveness of T4 for people under 65 with symptomatic subclinical hypothyroidism?

Why this is important:

Subclinical hypothyroidism (SCH) is a common biochemical abnormality that affects around 1% of people less than 70 years of age, rising to 6% in people in 80s. It is frequently transient and most people are asymptomatic. Large observational population surveys show that SCH is associated with increased vascular events, heart failure and mortality in younger individuals (50-70 age range). This may be because of a combination of dyslipidaemia caused by mild hypothyroidism and direct deficiency of thyroid hormone action on the myocardium. No randomised study of sufficient follow-up has been carried out that addressed the issue of long-term health outcomes in SCH. One high profile, randomised controlled trial studying the effect of subtherapeutic doses of levothyroxine (25 and 50mcg/d) in a largely asymptomatic group of older adults with SCH, on health-related QoL and symptoms at 12 months showed no clinically important benefit or harm from a low-dose of levothyroxine.

What remains unknown is whether symptomatic individuals with SCH aged <65 years could benefit from regular 'replacement' doses of levothyroxine both in terms of improvement in symptoms/QoL and of long-term cardiovascular events. Given the prevalence of SCH in women, it could be a major and entirely reversible cardiac risk factor but this idea remains essentially untested.

Criteria for selecting high-priority research recommendations:

PICO question	Population: Adults 50-70 years of age with persistent symptomatic SCH, defined as TSH >5.0mIU/l, normal FT4. Intervention(s): Levothyroxine ~1.0mcg/Kg/day; short duration-(?1yr) for QoL, 5yrs for major adverse cardiac events (MACE). Comparison: Placebo Outcome(s): symptoms, QoL (ThyPRO), MACE
Importance to patients or the population	A clinical trial would determine the effectiveness of treatment with therapeutic doses of levothyroxine for individuals aged 50-70 years. This could lead to improved treatment outcomes while minimising long-term cardiovascular risks and mortality at a very small cost.
Relevance to NICE guidance	This will address the lack of sufficient evidence to guide the management of symptomatic people with SCH using levothyroxine.
Relevance to the NHS	Evidence of the effect of levothyroxine in people with SCH younger than 65 / 50-70 years of age would ensure clinically and cost effective treatment for those people.
National priorities	There is a potential to improve CV event rates and mortality, which are national priorities
Current evidence base	Current evidence-base is limited to a single high profile study of limited follow-up which used a low dose of levothyroxine and looked for

	improvement in symptoms in a largely asymptomatic population
Equality	This disease is over-represented in women
Study design	Randomised, double blind placebo-controlled study of levothyroxine ~1.0 mcg/kg/d in people aged 50-70 years with persistent SCH for 3 months, 5 years intervention with MACE as the primary outcome
Feasibility	This would be an expensive study but a national 'randomisation, mail-out tablets for 5yrs, body count' approach through NHSCR would be feasible
Other comments	Our experience is that GPs frequently do measure TFTs and the current evidence-base leads to uncertainty and heterogeneous practice. This was suggested to NIHR back in ~2009 when BTA was asked for research suggestions but a multicentre application led by Cardiff was not funded largely due to costs >£6M
Importance	High. The guidelines are unable to provide clear recommendations for levothyroxine treatment for symptomatic people with SCH due to a lack of sufficient evidence. A well-executed randomised study could inform future updates.

K.2 Research question: What is the clinical and cost effectiveness of selenium oriodine for people with subclinical hypothyroidism?

Why this is important:

Subclinical hypothyroidism (SCH) is a biochemical abnormality that affects around 1% of people less than 70 years of age and 6% of people at 80 years. It is frequently transient and most people are asymptomatic. Large observational population surveys show that SCH is associated with increased vascular events, heart failure and mortality in individuals aged 50-70 years. This may be because of a combination of dyslipidaemia caused by mild hypothyroidism and direct deficiency of thyroid hormone action on the myocardium. SCH has furthermore been associated with increased all-cause and cardiovascular mortality.

Public interest regarding selenium andiodine supplementation for SCH was expressed at the scoping stage of this guideline. The metabolism of thyroid hormones requires iodine as key component but is also influenced by micronutrients including selenium. Existing studies have examined the effect of micronutrient supplementation on iodine status and the concentration of thyroid hormones, and observational evidence suggests that the concentration of micronutrients such as selenium is positively associated with iodine status. However, data from randomized controlled trials have failed to confirm this relationship.

Within the development of the present guideline, no evidence supporting selenium oriodine as treatment modalities for SCH was identified. There remains uncertainty regarding the efficacy and effectiveness of any of these treatments, although public interest remains high. There is therefore a need for a high quality trial to examine their clinical and cost effectiveness for the treatment of SCH.

Criteria for selecting high-priority research recommendations:

PICO question	Population: People with subclinical hypothyroidism Intervention(s):selenium and iodine Comparison: treatment with levothyroxine/ no treatment (placebo) Outcome(s): mortality, quality of life, adverse events (including

	cardiovascular, osteoporosis, impaired cognitive function) depression, hypothyroid/hyperthyroid symptoms, growth, TSH suppression
Importance to patients or the population	If any of these treatments provides clinically important benefits for people with SCH at a reasonable cost then it may be an important therapeutic modality to enhance clinical outcomes in this patient group that remains largely untreated.
Relevance to NICE guidance	This will address the lack of evidence to guide the management of people with SCH using selenium or iodine and guide future guideline development.
Relevance to the NHS	Evidence of the clinical and cost effectiveness of selenium and iodine for SCH would support the delivery of those as medicines for SCH ensuring the effective treatment of people with this condition.
National priorities	Considering the association of SCH with increased CV events, heart failure and mortality, identifying the effectiveness of the examined treatments provides the potential to improve CV event rates and mortality which are national priorities.
Current evidence base	There are currently no high quality studies supporting the effectiveness of selenium, or iodine in treating people with SCH to support the development of recommendations for their use.
Equality	There are no equality issues
Study design	Randomised, double blind placebo-controlled trial of selenium and iodine supplementation (3 months minimum duration) with long-term follow-up in people with SCH and a corresponding health economic analysis.
Feasibility	High interest of patients with SCH in the effectiveness of those treatments is likely to ensure the identification an adequate sample to enable the study.
Other comments	Both selenium and iodine are widely available on the market in the form of dietary supplements and can be obtained without prescription.
Importance	Low. The research is of interest to patients with SCH and will target the existing lack of evidence for selenium andiodinefor the treatment of SCH. However existing evidence suggests treatment of SCH in general, with conventional treatments like levothyroxine, does not result in clinically important benefits for most people.

1

2