National Institute for Health and Care Excellence

FINAL

Thyroid disease: assessment and management

[J] Management of thyrotoxicosis: anti thyroid drugs

NICE guideline NG145

Intervention evidence review underpinning recommendations 1.6.21 to 1.6.26 in the guideline. See also evidence reviews I, K, L and D

2019

FINAL

Developed by the National Guideline Centre, hosted by the Royal College of Physicians

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 All NICE guidance is subject to regular review and may be updated or withdrawn.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: 978-1-4731-3595-6

Contents

1 Management of thyrotoxicosis: pharmacological options 6

1.1 Review question: When anti-thyroid drugs are used, what is the most clinically and cost-effective way of using these drugs to treat thyrotoxicosis (for example choice of drugs, different treatment regimens)? 6

1.2 Introduction 6

1.3 PICO table 6

1.4 Clinical evidence 7

1.4.1 Included studies 7

1.4.2 Excluded studies 7

1.4.3 Summary of clinical studies included in the evidence review 8

1.4.4 Quality assessment of clinical studies included in the evidence review 14

1.5 Economic evidence 17

1.5.1 Included studies 17

1.5.2 Excluded studies 17

1.5.3 Health economic modelling 17

1.5.4 Resource costs 17

1.6 Evidence statements 17

1.6.1 Clinical evidence statements 17

1.6.2 Health economic evidence statements 18

1.7 The committee's discussion of the evidence 18

1.7.1 Interpreting the evidence 18

1.7.2 Cost effectiveness and resource use 21

1.7.3 Other factors the committee took into account 21

References 23

Appendices 31

Appendix A: Review protocols 31

Appendix B: Literature search strategies 37

Appendix C: Clinical evidence selection 46

Appendix D: Clinical evidence tables 47

Appendix E: Forest plots 77

E.2 Grave's disease- 12 – 18-month vs >18-month treatment 77

E.3 Grave's disease- 6 - <12-month vs 12 – 18-month treatment 78

E.4 Grave's disease- block-replace versus titration 78

Appendix F: GRADE tables 80

Appendix G: Health economic evidence selection 83

Appendix H: Health economic evidence tables 84

Appendix I: Health economic analysis 85

Appendix J: Excluded studies 86

Appendix K: Research recommendations 88

1

1 Management of thyrotoxicosis: pharmacological options

1.1 Review question: When anti-thyroid drugs are used, what is the most clinically and cost-effective way of using these drugs to treat thyrotoxicosis (for example choice of drugs, different treatment regimens)?

1.2 Introduction

Antithyroid drugs belong to the class of thionamides and have been in use to treat thyrotoxicosis since the 1940s. They inhibit the action of thyroperoxidase which is involved in the synthesis of thyroid hormones thereby blocking the thyroid gland and also may have direct and indirect effects on the immune system. The two main drugs used in the UK are carbimazole (its active component being methimazole) and propylthiouracil. Most patients with thyrotoxicosis associated with hyperthyroidism are started on antithyroid drugs to get control of the disease and in patients with Graves' disease a prolonged course (usually a minimum of 12 months) may be given to try and induce remission of the disease.

Both carbimazole and propylthiouracil may have minor side effects in up to 5% of people using these medications. These adverse effects include cutaneous allergic reactions, arthralgias and gastro-intestinal upset. Severe side-effects are rare but include agranulocytosis in 0.2-0.5% of patients and may occur with both drugs. Propylthiouracil has been linked to hepatotoxicity and vasculitis and is therefore less commonly used first line, whereas carbimazole is associated with an increased risk of teratogenicity and pancreatitis.

Antithyroid drugs may be used in a titration regimen where a relatively high dose is started and the dose is gradually reduced over the next weeks to months depending on the response to treatment. Alternatively, a block and replace regime may be used where high doses of antithyroid drugs are continued and levothyroxine is added to maintain biochemical euthyroidism.

This review will focus on the optimal duration of a course of antithyroid drugs, the choice of medication and the regimen to be used in order to achieve lasting remission of thyrotoxicosis and to avoid adverse events.

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 thyrotoxicosis (TSH below normal reference ranges, free T3/T4 above normal reference range)
Interventions	• Carbimazole/methimazole vs propylthiouracil • Block and replace vs titration regimen • 6-<12 months vs 12-18 months vs >18 months
Comparisons	Comparisons between modalities
Outcomes	**Critical** • Mortality (dichotomous, ≥1 year) • Quality of life (continuous) **Important**

	• Thyroid ophthalmopathy (dichotomous) • Euthyroidism (dichotomous) • Hypothyroidism (dichotomous) • Relapse of hyperthyroidism (dichotomous) • Cardiovascular morbidity (ischaemic heart disease, dichotomous) • Arrhythmia (dichotomous) • Osteoporosis (dichotomous) • Cognitive impairment (dichotomous) • Pain (continuous) • Symptom scores (continuous) • Patient/family/carer experience (continuous) • Healthcare contacts (rates/dichotomous) • Agranulocytosis (dichotomous) • Liver failure (dichotomous) • Minor drug related adverse effects (dichotomous) • Teratogenesis (dichotomous)
Study design	• Minimum follow-up of 3 months • RCTs • Non-randomised cohort studies to be considered if adjusted for key confounders (age, co-existing conditions, baseline T4, size of goitre) and insufficient RCTs evidence found, on an intervention by intervention basis

1.4 Clinical evidence

1.4.1 Included studies

Fifteen studies were included in the review; [3, 27, 29, 32, 35, 37, 52, 58, 59, 63, 65, 68, 77, 79, 91] these are summarised in Table 2 below. Evidence from these studies is summarised in the clinical evidence summary below (Table 3). One Cochrane review in this area was identified[2], the studies included in this review were checked against the protocol and included as appropriate.

Four studies were found comparing methimazole or carbimazole with propylthiouracil in people with Graves' disease. [35, 37, 63, 68]

Four studies were found comparing the efficacy of different treatment durations (long vs short-term) with antithyroid drugs (carbimazole) for Graves' disease. Two of those studies compared a 12-18 month treatment with treatment exceeding 18 months. [29, 58] The remaining two studies compared a 6-<12 month treatment with a 12-18 month treatment. [3, 91]

Seven studies compared a block and replace treatment regimen with a titrated treatment regimen. [27, 32, 52, 59, 65, 77, 79]

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
Allanic 1990 [3]	12-18m (18-month treatment): carbimazole 30-60 mg/d, later reduced to 10-20 mg/d to maintain euthyroidism, n=57 6-<12 m (6-month treatment): carbimazole 30-60 mg/d, later reduced to 10-20 mg/d to maintain euthyroidism, n=57	Adults (18 month mean age 39.2 SD 12.3; 6-month mean age 43.1 SD 14.7) Graves' disease Treatment naïve France	Euthyroidism Relapse 24 months after treatment withdrawal	Parallel design Titration
Edmonds 1994 [27]	Block-replace: carbimazole 60mg/d; T4 100-150 µg/d (beginning four weeks later), 12 months, n=49 Titration: carbimazole 60mg/d for four weeks, then reduced to reach maintenance dose (usually by the third month of treatment), 12 months, n=46	Adults (block-replace mean age 48 SD 11.9; titration mean age 41, SD 12.9) Graves' disease Treatment naïve United Kingdom	Relapse Minor drug related adverse events (during treatment) Agranulocytosis (during treatment) 24 months after treatment withdrawal	Parallel design
García-Mayor 1992 [29]	> 18 m (24-month treatment): Carbimazole 10 mg every 8 h; reduced to ≥10 mg/day to maintain euthyroidism once reached for 24 months, n=24 12-18m (12-month treatment): Carbimazole 10 mg every 8 h; reduced to ≥10 mg/day to maintain euthyroidism once	Adults (mean age 39.35, SD 13.69) Graves' disease Treatment naïve Spain	Relapse 5 years after treatment withdrawal	Parallel design Titration

Study	Intervention and comparison	Population	Outcomes	Comments
	reached for 12 months , n=28			
Grebe 1998 [32]	Block-replace: carbimazole 100 mg/d; T4 starting 2-3 weeks later, adjusted to maintain euthyroidism, mean dose increased to 132µg/d at end of treatment, n=17 Titration: 25 mg/d titrated to maintain euthyroidism (average 17 mg/d at end of treatment), n=20	Adults (block-replace mean age 33 SD 8.7; titration mean age 33.7 SD 11.9) Graves' disease Newly diagnosed New Zealand	Relapse (24 months after treatment withdrawal) Minor drug related adverse events Agranulocytosis During the 6-month treatment	Parallel design
He 2004 [35]	MMI: 15 mg/d, n=15 PTU: 150 mg/d, n=15	Adults (MMI mean age 32, SD 7.1; PTU mean 31 SD 6.5) Graves' disease Newly diagnosed China	Euthyroidism Hypothyroidism 12 weeks	Parallel design Stable dose
Homsanit 2001 [37]	MMI: 15 mg/d, n=35 PTU: 150 mg/d, n=36	Adults (MMI mean age 35.4, SD 11.5; PTU mean 34.8 , SD 13) Graves' disease Newly diagnosed Thailand	Euthyroidism Hypothyroidism 12 weeks	Parallel design Stable dose
Lucas 1997[52]	Both initially received carbimazole 45-60 mg/d until achievement of euthyroidism	Adults (block-replace mean age 34.5 SD 8.3; titration mean age 37.5 SD 13.9)	Relapse Mean (SD) 8.5 (8.7) months	Parallel design

Study	Intervention and comparison	Population	Outcomes	Comments
	Block-replace: carbimazole 30-45 mg/d, T4 100 µg/d adjusted after 1 months to 75-150 mg/d to maintain euthyroidism (normal FT4, T3), n=30 Titration: carbimazole doses adjusted to maintain euthyroidism, n=30	Graves' disease Treatment naïve Spain	after carbimazole treatment withdrawal; overall mean (SD) follow up 4.98(1.6) years.	
Maugendre 1999 [58]	> 18 m (42-month treatment): Carbimazole 20-50 mg/d for 3 months followed by 10-15 mg/d to maintain euthyroidism for 42 months, n=62 12-18m (18-month treatment): Carbimazole 20-50 mg/d for 3 months followed by 10-15 mg/d to maintain euthyroidism for 18 months, n=72	Adults (median age (range): 40.4 (13-74) Graves' disease Treatment naïve France	Relapse 24 months after treatment withdrawal	Parallel design Titration
McIver 1996 [59]	Block-replace: carbimazole 20 mg twice daily; T4 100mg/d initially, adjusted to achieve undetectable TSH (<0.04 µU/ml), for 17 months (plus T4 alone continued for 18 months), n=59 Titration: carbimazole 40mg/d, adjusted to achieve normal TSH, T4 and T3 for 17 months,n=52	Adults (block-replace mean age 36 SD 10; titration mean age 33 SD 9) Graves' disease Treatment naïve United Kingdom	Relapse Median 12 months (median 3-18 months) after treatment withdrawal	Parallel design

Study	Intervention and comparison	Population	Outcomes	Comments
	Both groups initially given carbimazole 40mg/d for one month			
Nakamura 2007 [63]	MMI: 15 (single dose) or 30 (two divided doses) µg/d , reduced to 10 mg or 15 mg respectively, n=282 PTU: 300 µg/d (three divided doses), reduced to 150 mg, n=114	Adults (MMI mean age 20.29, SD 13.3; PTU mean age 40.2, SD 12.9) Graves' disease Newly diagnosed/ treatment naive Japan	Euthyroidism Drug related adverse events 12 weeks	Multicentre study Parallel design Titration
Nedrebo 2002 [65]	Block-replace: carbimazole dose mean (range) 29.7 mg/d (15-45 mg) except for one patient receiving propylthiouracil 200-400 mg/d; L-T4 to maintain normal FT-4 once euthyroid, n=110 Titration: carbimazole dose mean (range) 29.7 mg/d (15-45 mg) except for five patients receiving propylthiouracil 200-400 mg/d; initial dose adjusted to maintain normal FT-4 once euthyroid, n=108	Adults: (block-replace mean age 42.02, SD 11.04; titration mean age 42.8, SD 12.77) Graves' disease No previous treatment with ATD drugs for at least 12 months Norway	Relapse 24 months after ATD withdrawal	Parallel design
Peixoto 2006 [68]	MMI: 40 to 60 mg daily, n=30 PTU: 200 to 300 mg every 12 hours, n=25	Adults (mean age 37.7, SD 10.5) Graves' disease	Euthyroidism Minor drug related adverse events	Parallel design Titration

Study	Intervention and comparison	Population	Outcomes	Comments
		Treatment naïve Brazil	12-month treatment; 12-38 month follow-up	
Rittmaster 1998[77]	Block-replace: 15 mg MMI twice daily for 18 months & T4 sufficient dose to maintain TSH in the mid- to high-normal range: 2.0-5.4 mIU/L (n=50) or TSH less than or equal to 0.6 mIU/L (n=48), n=98 Titration: MMI for 18 months, adjusted to maintain normal TSH (0.3-5.4 mIU/L), n=51 Both groups initially treated with MMI 10 mg three times daily for mean (SD) 7.9(6.2) weeks until normal T3 (0.9-2.8 nmol/L) reached.	Adults (mean age 38 SD 14) Graves' disease Treatment naïve Canada	Relapse Mean follow up 27 months (Range 6-47) after treatment withdrawal	Parallel design
Romaldini 1983[79]	Block-replace: MMI 40-100 mg/d (mean (SD) 60.7 (14.5) n=34) or PTU 500-1200 mg/d (mean (SD) 694 (173), n=31); large start dose, increased to obtain total blockage when necessary; 50-75 µg T3 added 2-3 weeks after, n=65 10-30 month treatment, mean (SD): 15.1 (4.2)	Adults (block-replace mean age 40, SD 11; titration mean age 40, SD 13) Graves' disease Brazil	Relapse After treatment withdrawal mean (SD) 42(14) months, range: 17-81 months	Parallel design

Study	Intervention and comparison	Population	Outcomes	Comments
	Titration: MMI 40 mg or PTU 500 mg, gradually reduced to MMI 5-25 mg/d (mean (SD):13.6(7), n=25) or PTU 100-300 mg/d (mean (SD): 180(58), n=23) to maintain euthyroid state, 12-20 months n=48 12-20 month treatment, mean(SD): 13.5 (2.2)			
Weetman 1994[91]	12-18m (12 months): carbimazole 20mg three times/day, reduced to 40 mg single dose after 4 weeks; thyroxin started at 4 weeks at 1.5 mcg/kg daily, rounded up to the nearest 25 mcg if the patient was euthyroid or deferred for 1-2 weeks if patient was still hyperthyroid, n=51 6-<12m (6 months): carbimazole 20mg three times/day, reduced to 40 mg single dose after 4 weeks; thyroxin started at 4 weeks at 1.5 mcg/kg daily, rounded up to the nearest 25 mcg if the patient was euthyroid or deferred for 1-2 weeks if patient was still hyperthyroid, n=49	Adults >55 Graves' disease Treatment naïve United Kingdom	Euthyroidism 12 months after treatment withdrawal	Parallel design Block-replace treatment

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: Methimazole/carbimazole versus propylthiouracil

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Propylthiouracil	Risk difference with Methimazole/carbimazole (95% CI)
Euthyroidism cases	410 (4 studies) 3-12 months	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, inconsistency, imprecision	RR 1.51 (0.75 to 3.03)	524 per 1000	267 more per 1000 (from 131 fewer to 1000 more)
Hypothyroidism cases	101 (2 studies) 12 weeks	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision	Peto OR 10.27 (3.46 to 30.44)	0 per 1000	300 more per 1000 (from 160 more to 440 more)[4]
Minor drug related adverse events	417 (2 studies) 3-12 months	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, inconsistency	RR 0.44 (0.33 to 0.59)	260 per 1000	146 fewer per 1000 (from 107 fewer to 174 fewer)

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 or 2 increments because the point estimate and or the confidence intervals varied widely across studies, unexplained by subgroup analysis
3 Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs
4 Zero events in control group

Table 4: Clinical evidence summary: 12 - 18 month versus >18 month treatment

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with 12-18m	Risk difference with >18m (95% CI)
Relapse cases (post treatment withdrawal)	186 (2 studies)	⊕⊖⊖⊖ VERY LOW1, 2	RR 0.88 (0.67 to 1.16)	609 per 1000	73 fewer per 1000 (from 201 fewer to 97 more)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with 12-18m	Risk difference with >18m (95% CI)
	2-5 years	due to risk of bias, imprecision			

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

Table 5: Clinical evidence summary: 6- <12 month versus 12-18 month treatment

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with 6-12m	Risk difference with 12-18m (95% CI)
Relapse cases (post treatment withdrawal)	94 (1 study) 24 months	⨁⨁⨁⊖ MODERATE1 due to imprecision	RR 0.63 (0.41 to 0.99)	583 per 1000	216 fewer per 1000 (from 6 fewer to 344 fewer)
Euthyroidism cases (post treatment withdrawal)	194 (2 studies) 12-24 months	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision	RR 1.26 (0.99 to 1.61)	504 per 1000	131 more per 1000 (from 5 fewer to 307 more)

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

Table 6: Clinical evidence summary: Block-replace versus titration

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Titration	Risk difference with Block-replace (95% CI)
Relapse	659	⨁⊖⊖⊖	RR 0.8	583 per	117 fewer per 1000

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Titration	Risk difference with Block-replace (95% CI)
cases	(7 studies) 6-47 months	VERY LOW1,2,3 due to risk of bias, inconsistency, imprecision	(0.63 to 1.03)	1000	(from 216 fewer to 17 more)
Minor drug related adverse events (during treatment)	132 (2 studies) 6-12 months	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, inconsistency, imprecision	RR 2.70 (0.23 to 31.79)	65 per 1000	111 more per 1000 (from 50 fewer to 935 more)
Agranulocytosis cases (during treatment)	132 (2 studies) 6-12 months	⨁⨁⊖⊖ LOW1,3 due to risk of bias, imprecision	Peto OR 3.19 (0.43 to 23.74)	25 per 1000	51 more per 1000 (from 14 fewer to 353 more)

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 or 2 increments because the point estimate and or the confidence interval varied widely across studies, unexplained by subgroup analysis.
3 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 Anti-thyroid drugs

Drug	Daily dose	Cost – month	Cost – annual
Propylthiouracil 50mg tablets	50mg to 150mg (a)	£13 - £40	£160 - £480
Carbimazole 5mg tablets	5mg to 15mg (b)	£6.50 - £19	£78 - £234
Carbimazole 20mg +Levothyroxine 100µg tablets	40mg + 100µg (c)	£57 (d)	£685 (e)

Source: BNF, Date, August 2019.[141414](BMJ Group and the Royal Pharmaceutical Society of Great Britain)
(a) Maintenance dose reported, initial dose of 200-400 mg daily in divided doses until patient becomes euthyroid.
(b) Maintenance dose reported based on 5mg tablets, initial dose of 15-40 mg daily in divided doses until patient becomes euthyroid, usually 4 to 8 weeks.
(c) Blocking- replacement regimen; combination of carbimazole 40- 60 mg daily, with levothyroxine 100µg daily, usually given for 18 months.
*(d) Rounded up cost of carbimazole 40mg (2 * 20mg tablets) and levothyroxine 100µg; £55.97 + £1.03*
*(e) Rounded up cost of carbimazole 40mg (2 * 20mg tablets) and levothyroxine 100µg; £671.67 + £13.43*

1.6 Evidence statements

1.6.1 Clinical evidence statements

1.6.1.1 Methimazole/carbimazole vs propylthiouracil

There was a clinically important benefit of methimazole/carbimazole for euthyroidism (4 studies, Very Low quality) and minor drug related adverse events (2 studies, Very Low quality).

There was a clinically important harm of methimazole/carbimazole for hypothyroidism (2 studies, Very Low quality).

No evidence was identified for other outcomes.

1.6.1.2 12-18 month vs >18 month treatment

No clinically important difference was identified for relapse to hyperthyroidism (2 studies, Very Low quality).

No evidence was identified for other outcomes.

1.6.1.3 12-18 month treatment vs 6-<12 month

There was a clinically important benefit of a 12-18 month treatment for relapse (1 study, Moderate quality) and euthyroidism (2 studies, Low quality).

No evidence was identified for other outcomes.

1.6.1.4 Block and replace vs titration

There was a clinically important benefit of a block and replace treatment regimen for relapse (7 studies, Very Low quality).

There was a clinically important harm of a block and replace treatment regimen for minor drug related adverse events (2 studies, Very Low quality) agranulocytosis (2 studies, Low quality).

No evidence was identified for other outcomes.

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 included thyroid ophthalmopathy, euthyroidism, hypothyroidism, relapse of hyperthyroidism, cardiovascular morbidity, arrhythmia, osteoporosis, cognitive impairment, pain, symptom scores, experience of care, healthcare contacts, agranulocytosis, liver failure, minor drug related adverse effects, teratogenesis.

Considering there was no clinical evidence in regard to the critical outcomes, it was agreed that decision making would be based on the important outcomes of this review for which there was evidence.

1.7.1.2 The quality of the evidence

The quality of the evidence in this review ranged from moderate to very low quality, with the majority being very low quality. Evidence was typically downgraded for risk of bias often attributed to lack of blinding and methodological shortcomings such as differences in the length of follow-up between different groups and imprecision. Some comparisons were downgraded for inconsistency that could not be explained by any subgroup analysis pre-specified in the review protocol.

No evidence was identified for any comparison in children or older adults. No evidence was identified in people with thyrotoxicosis with a diagnosis other than Graves' disease or who have had previous treatment for thyrotoxicosis.

The committee noted that the doses of carbimazole (60mg a day and 100mg a day) used in some studies were considerably higher than what is currently being used in the UK. Lower doses would be expected to result in less hypothyroidism and potentially more euthyroidism than what was identified in the case of different antithyroid drug comparisons. More commonly seen lower doses (e.g. 30-40mg) were considered to be more appropriate and could lead to fewer adverse events than those identified in this review, which the committee agreed could be linked to higher dose regimens.

Relapse of hyperthyroidism after treatment withdrawal was the most frequently reported outcome. Evidence for the majority of outcomes included in the protocol by the committee, including the critical outcomes of mortality and quality of life was not identified.

The committee noted that there was very little evidence about rare but serious and well-established adverse events of drugs (e.g. teratogenesis, severe liver damage). This was unsurprising given the length of follow-up and number of participants.

All evidence in the review related to the use of antithyroid drugs as definitive treatment for thyrotoxicosis and the majority of the discussion below pertains to this scenario. However the committee noted that some of the evidence would still apply to when antithyroid drugs were used short term in advance of other definitive treatment (for example in selecting which drugs would have fewer side effects).

1.7.1.2.1 Methimazole/carbimazole vs propylthiouracil

The quality of the evidence regarding the use of methimazole/carbimazole compared to propylthiouracil was very low and was downgraded for risk of bias. The evidence was furthermore downgraded for inconsistency unexplained by protocol pre-specified subgroup analysis and imprecision. Studies relative to this antithyroid drug comparison had a short-term follow-up period with participants being followed for up to 12 months.

1.7.1.2.2 6-<12 months vs 12-18 months vs >18 months

The quality of the evidence relative to the 12-18 month and the 18-month treatment comparison was very low and was downgraded due to risk of bias and imprecision. Participants within this comparison were followed for up to 5 years post treatment withdrawal.

The quality of the evidence relative to the 6 -<12 month and the 12-18 month treatment comparison ranged from moderate to low and was downgraded for imprecision. One outcome was also downgraded due to risk of bias. Participants within this comparison were followed for up to 24 months after treatment withdrawal.

1.7.1.2.3 Block and replace vs titration

Within the block and replace and titration treatment regimen comparison the majority of the evidence was of very low quality with evidence for one outcome (agranulocytosis) being of low quality. The evidence was generally downgraded for risk of bias and imprecision. Two comparisons were also downgraded for inconsistency that could not be explained by subgroup analyses. The studies included in this comparison had relatively short follow-up periods with the majority of participants followed up for 12-24 months post treatment withdrawal.

1.7.1.3 Benefits and harms

1.7.1.3.1 Methimazole/carbimazole vs propylthiouracil

The evidence showed that methimazole/carbimazole has a clinically important benefit compared with propylthiuracil in terms of euthyroidism and the emergence of minor drug related adverse events such as the development of skin rash.

Compared with propylthiuracil, methimazole/carbimazole also appeared to lead to more people ending up at a hypothyroid state. The committee discussed the outcome of hypothyroidism noting it would be unlikely to constitute a permanent outcome in people treated for Graves' disease with antithyroid drugs. There was agreement that hypothyroidism was likely to be the result of over treating that would involve using a higher than appropriate dose or failure to reduce the initial dose when appropriate. Considering the shortterm follow-up of the included studies (12 months), within this drug comparison hypothyroidism was not considered to be a meaningful outcome in decision making.

The committee agreed that it would not be appropriate to use propylthiuracil in children given its association with severe liver damage; this is a well-established adverse event although not captured in this evidence review.

The committee emphasised that while carbimazole should be first line, propylthiuracil may have use in pregnancy/planned pregnancy, in people who cannot tolerate carbimazole or are particularly likely to suffer adverse events (for example those with a history of pancreatitis) .

1.7.1.3.2 6-<12 months vs 12-18 months vs >18 months

Compared to 18-month treatments with antithyroid drugs (carbimazole), longer treatments exceeding 18 months (24 and 42-month treatments) did not appear to lead to a clinically important difference in terms of relapse to hyperthyroidism two to five years after treatment withdrawal.

Compared to 6-month treatments with carbimazole, 12 and 18 month treatments led to a clinically important benefit both in terms of relapse to hyperthyroidism and euthyroidism (1-2 years) after treatment withdrawal.

Therefore overall, treating for 12-18 months appeared to have a benefit over shorter treatment periods but there was no benefit of treating for longer than 18 months. The committee also noted that in some situations more prolonged courses could have benefits, for example in children (though regular review would be required) or in those with thyrotoxicosis secondary to a cause that will not resolve like multiple toxic nodules that cannot be treated with surgery or radioactive iodine.

The committee noted that at the time of stopping other clinical factors will affect the decision as to whether to stop, for instance the antibody status of the person at the time of stopping, their TSH and the dose of antithyroid drugs required to maintain euthyroidism.

The committee noted that the purpose of antithyroid drug treatment in this context is to control hyperthyroidism until the underlying autoimmune process resolves spontaneously. The aim of treatment is not merely to treat until euthyroidism and then stop treatment.

1.7.1.3.3 Block and replace vs titration

The evidence showed that block and replace treatment regimens have a clinically important benefit compared with titration regimens in terms of relapse to hyperthyroidism. The committee noted that this constitutes an interesting finding considering that no such difference has been previously documented in the evidence existing to date. However, the committee also noted the very low-quality evidence underpinning this finding.

There was a clinically important harm of block and replace in minor drug related adverse events including skin reactions, itchiness and skin rash compared to titration treatments identified in this review although the evidence was very low quality with very serious imprecision.

The committee noted that there was a clinically important harm for block and replace regimens in terms of agranulocytosis. This a serious adverse event which, in the committee's experience, can lead to death in approximately 10% of cases. However, the evidence for this comparison was based on very low event rates. Furthermore, the carbimazole doses were higher than what is seen in current practice which may have contributed to this effect.

The committee noted a number of characteristics that may help inform the choice to use a block and replace or titration regimen. A block and replace regimen is an option when a stable thyroid status is desirable at an early stage (e.g. young people with examinations pending or when there is established thyroid eye disease).

1.7.2 Cost effectiveness and resource use

There was no health economic evidence identified for this review question. The Committee considered the costs of the different drugs alongside the clinical evidence to make a judgement regarding likely cost effectiveness.

Carbimazole was found to be lower cost than propylthiuracil and the committee concluded it was more clinically effective; they therefore concluded it was cost effective.

The committee noted that treating the patient and reviewing their progress at 12-18 months, which was found to be clinically beneficial and current practice, can also be cost saving, compared to shorter (<12) or longer (>18 months) treatment. This is because not all patients would need further drugs hence avoiding unnecessary prescribing. In addition, patients that may need dose adjustments or alternative interventions can be identified earlier which may lead to less complications, better quality of life and reduced downstream costs. The committee noted that the review would be carried out at an existing appointment and therefore unlikely to incur additional costs.

The cost of block and replace regimen was found to be £685 per year based on 40mg carbimazole and 100µg of levothyroxine compared to the titration regimen based on carbimazole 5mg tablets being titrated between 5mg to 15mg costing £78 to £234 per year. The committee noted that the cost of the block and replace regime may be offset by a reduction in hospital visits and number of blood tests required. However, given no clinical evidence was identified in this area cost effectiveness is uncertain. The committee chose therefore to make recommendations to make physicians aware when choosing between titration, or block and replace regimens, to consider the factors listed here, to target treatment to those that would be most likely to benefit from treatment based on their clinical experience. The committee also made a research recommendation to help assess the cost effectiveness between the two regimes.

In children with Graves' disease, titration doses are currently used, as they are associated with fewer side effects. Propylthiouracil is not recommended in children due to its hepatotoxic effects. Toxic nodular goitre is rare in children, and current practice is to remove it via surgery which is unlikely to have cost implications.

1.7.3 Other factors the committee took into account

In adults, people tend to opt to switch to radioactive iodine or surgery after a failed course of antithyroid drugs. Children and their families are more reluctant to opt for radioactive iodine/surgery.

Typically, if relapse occurs after successful antithyroid drug treatment, people tend to restart their previous antithyroid drug regimen (assuming they do not opt to switch to surgical or radioactive iodine treatment). People do not commonly opt for an additional antithyroid drug treatment but in an alternative form (for example switching from carbimazole to propylthiouracil or switching from block and replace to titration regimen), although use of titration regimens in older people opting against any definitive treatment is not uncommon.

Current practice in the UK, in adults and children, is a mix of block and replace (~40%) and titration regimens (~60%). It is generally agreed that theoretically block and replace regimens require less follow-up and monitoring, although this is not definitive.

References

1. Abraham-Nordling M, Wallin G, Lundell G, Törring O. Thyroid hormone state and quality of life at long-term follow-up after randomized treatment of Graves' disease. European Journal of Endocrinology. 2007; 156(2):173-179

2. Abraham P, Avenell A, McGeoch SC, Clark LF, Bevan JS. Antithyroid drug regimen for treating Graves' hyperthyroidism. Cochrane Database of Systematic Reviews 2010, Issue 1. Art. No.: CD003420. DOI: 10.1002/14651858.CD003420.pub4.

3. Allannic H, Fauchet R, Orgiazzi J, Madec AM, Genetet B, Lorcy Y et al. Antithyroid drugs and Graves' disease: a prospective randomized evaluation of the efficacy of treatment duration. Journal of Clinical Endocrinology and Metabolism. 1990; 70(3):675-679

4. Andrade VA, Gross JL, Maia AL. Effect of methimazole pretreatment on serum thyroid hormone levels after radioactive treatment in Graves' hyperthyroidism. Journal of Clinical Endocrinology and Metabolism. 1999; 84(11):4012-4016

5. Andrade VA, Gross JL, Maia AL. The effect of methimazole pretreatment on the efficacy of radioactive iodine therapy in Graves' hyperthyroidism: one-year follow-up of a prospective, randomized study. Journal of Clinical Endocrinology and Metabolism. 2001; 86(8):3488-3493

6. Andrade VA, Gross JL, Maia AL. Serum thyrotropin-receptor autoantibodies levels after I therapy in Graves' patients: effect of pretreatment with methimazole evaluated by a prospective, randomized study. European Journal of Endocrinology. 2004; 151(4):467-474

7. Azizi F, Amouzegar A. Management of thyrotoxicosis in children and adolescents: 35 years' experience in 304 patients. Journal of Pediatric Endocrinology and Metabolism. 2018; 31(2):159-165

8. Azizi F, Yousefi V, Bahrainian A, Sheikholeslami F, Tohidi M, Mehrabi Y. Long-term continuous methimazole or radioiodine treatment for hyperthyroidism. Archives of Iranian Medicine. 2012; 15(8):477-484

9. Barczy?ski M, Konturek A, Hubalewska-Dydejczyk A, Go?kowski F, Nowak W. Randomized clinical trial of bilateral subtotal thyroidectomy versus total thyroidectomy for Graves' disease with a 5-year follow-up. British Journal of Surgery. 2012; 99(4):515-522

10. Barczynski M, Konturek A, Hubalewska-Dydejczyk A, Golkowski F, Miklaszewska G, Romanowska-Dixon B. Five-year follow up of a randomized clinical trial of bilateral subtotal thyroidectomy versus total thyroidectomy for Graves' disease. Langenbeck's Archives of Surgery. 2010; 395(4):471

11. Barczynski M, Konturek A, Hubalewska-Dydejczyk A, Golkowski F, Nowak W. Ten-year follow-up of a randomized clinical trial of total thyroidectomy versus Dunhill operation versus bilateral subtotal thyroidectomy for multinodular non-toxic goiter. World Journal of Surgery. 2018; 42(2):384-392

12. Benker G, Reinwein D, Kahaly G, Tegler L, Alexander WD, Fassbinder J et al. Is there a methimazole dose effect on remission rate in Graves' disease? Results from a long-term prospective study. The European Multicentre Trial Group of the Treatment of Hyperthyroidism with Antithyroid Drugs. Clinical Endocrinology. 1998; 49(4):451-457

13. Benker G, Vitti P, Kahaly G, Raue F, Tegler L, Hirche H et al. Response to methimazole in Graves' disease. The European Multicenter Study Group. Clinical Endocrinology. 1995; 43(3):257-263

14. 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

15. Bonnema SJ, Bennedbaek FN, Gram J, Veje A, Marving J, Hegedus L. Resumption of methimazole after 131I therapy of hyperthyroid diseases: effect on thyroid function and volume evaluated by a randomized clinical trial. European Journal of Endocrinology. 2003; 149(6):485-492

16. Bonnema SJ, Bennedbaek FN, Veje A, Marving J, Hegedüs L. Propylthiouracil before 131I therapy of hyperthyroid diseases: effect on cure rate evaluated by a randomized clinical trial. Journal of Clinical Endocrinology and Metabolism. 2004; 89(9):4439-4444

17. Bonnema SJ, Grupe P, Boel-Jørgensen H, Brix TH, Hegedüs L. A randomized trial evaluating a block-replacement regimen during radioiodine therapy. European Journal of Clinical Investigation. 2011; 41(7):693-702

18. Braga M, Walpert N, Burch HB, Solomon BL, Cooper DS. The effect of methimazole on cure rates after radioiodine treatment for Graves' hyperthyroidism: a randomized clinical trial. Thyroid. 2002; 12(2):135-139

19. Burch HB, Solomon BL, Cooper DS, Ferguson P, Walpert N, Howard R. The effect of antithyroid drug pretreatment on acute changes in thyroid hormone levels after (131)I ablation for Graves' disease. Journal of Clinical Endocrinology and Metabolism. 2001; 86(7):3016-3021

20. Buscemi S, Verga S, Cottone S, Andronico G, D'Orio L, Mannino V et al. Favorable clinical heart and bone effects of anti-thyroid drug therapy in endogenous subclinical hyperthyroidism. Journal of Endocrinological Investigation. 2007; 30(3):230-235

21. Canto AU, Dominguez PN, Jimeno CA, Obaldo JM, Ogbac RV. Comparison of Fixed versus Calculated Activity of Radioiodine for the Treatment of Graves Disease in Adults. Endocrinol metab. 2016; 31(1):168-173

22. Chen DY, Schneider PF, Zhang XS, He ZM, Jing J, Chen TH. Striving for euthyroidism in radioiodine therapy of Graves' disease: a 12-year prospective, randomized, open-label blinded end point study. Thyroid. 2011; 21(6):647-654

23. Chen DY, Schneider PF, Zhang XS, Luo XY, He ZM, Chen TH. Changes in graves' ophthalmopathy after radioiodine and anti-thyroid drug treatment of graves' disease from 2 prospective, randomized, open-label, blinded end point studies. Experimental and clinical endocrinology & diabetes. 2014; 122(1):1-6

24. Chi SY, Hsei KC, Sheen-Chen SM, Chou FF. A prospective randomized comparison of bilateral subtotal thyroidectomy versus unilateral total and contralateral subtotal thyroidectomy for graves' disease. World Journal of Surgery. 2005; 29(2):160-163

25. Connell JM, Hilditch TE, Robertson J, Coghill G, Alexander WD. Radioprotective action of carbimazole in radioiodine therapy for thyrotoxicosis--influence of the drug on iodine kinetics. European Journal of Nuclear Medicine. 1987; 13(7):358-361

26. De Luca F, Valenzise M. Controversies in the pharmacological treatment of Graves' disease in children. Expert Review of Clinical Pharmacology. 2018:1-9

27. Edmonds CJ, Tellez M. Treatment of Graves' disease by carbimazole: high dose with thyroxine compared to titration dose. European Journal of Endocrinology. 1994; 131(2):120-124

28. Esfahani AF, Kakhki VR, Fallahi B, Eftekhari M, Beiki D, Saghari M et al. Comparative evaluation of two fixed doses of 185 and 370 MBq 131I, for the treatment of Graves' disease resistant to antithyroid drugs. Hellenic Journal of Nuclear Medicine. 2005; 8(3):158-161

29. García-Mayor RV, Páramo C, Luna Cano R, Pérez Mendez LF, Galofré JC, Andrade A. Antithyroid drug and Graves' hyperthyroidism. Significance of treatment duration and TRAb determination on lasting remission. Journal of Endocrinological Investigation. 1992; 15(11):815-820

30. Glinoer D, Nayer P, Bex M. Effects of l-thyroxine administration, TSH-receptor antibodies and smoking on the risk of recurrence in Graves' hyperthyroidism treated with antithyroid drugs: a double-blind prospective randomized study. European Journal of Endocrinology. 2001; 144(5):475-483

31. Goni Iriarte MJ, Forga Llenas L, Iriarte Beroiz A, Anda Apinaniz E, Rodriguez Erdozain R, Menendez Torre E. Recurrence of Graves' disease: the influence of treatment schedule. Medicina Clínica. 1995; 104(1):11-14

32. Grebe SK, Feek CM, Ford HC, Fagerström JN, Cordwell DP, Delahunt JW et al. A randomized trial of short-term treatment of Graves' disease with high-dose carbimazole plus thyroxine versus low-dose carbimazole. Clinical Endocrinology. 1998; 48(5):585-592

33. Hamide AKAHDSJS. Radioiodine therapy in patients with Graves' disease and the effects of prior carbimazole therapy. Indian Journal of Endocrinology and Metabolism. 2014; 18(5):688-693

34. Hashizume K, Ichikawa K, Sakurai A, Suzuki S, Takeda T, Kobayashi M et al. Administration of thyroxine in treated Graves' disease. Effects on the level of antibodies to thyroid-stimulating hormone receptors and on the risk of recurrence of hyperthyroidism. New England Journal of Medicine. 1991; 324(14):947-953

35. He CT, Hsieh AT, Pei D, Hung YJ, Wu LY, Yang TC et al. Comparison of single daily dose of methimazole and propylthiouracil in the treatment of Graves' hyperthyroidism. Clinical Endocrinology. 2004; 60(6):676-681

36. Hoermann R, Quadbeck B, Roggenbuck U, Szabolcs I, Pfeilschifter J, Meng W et al. Relapse of Graves' disease after successful outcome of antithyroid drug therapy: results of a prospective randomized study on the use of levothyroxine. Thyroid. 2002; 12(12):1119-1128

37. Homsanit M, Sriussadaporn S, Vannasaeng S, Peerapatdit T, Nitiyanant W, Vichayanrat A. Efficacy of single daily dosage of methimazole vs. propylthiouracil in the induction of euthyroidism. Clinical Endocrinology. 2001; 54(3):385-390

38. Howarth D, Epstein M, Lan L, Tan P, Booker J. Determination of the optimal minimum radioiodine dose in patients with Graves' disease: a clinical outcome study. European Journal of Nuclear Medicine. 2001; 28(10):1489-1495

39. Jaiswal AK, Bal C, Damle NA, Ballal S, Goswami R, Hari S et al. Comparison of clinical outcome after a fixed dose versus dosimetry-based radioiodine treatment of Graves' disease: Results of a randomized controlled trial in Indian population. Indian Journal of Endocrinology and Metabolism. 2014; 18(5):648-54

40. Järhult J, Rudberg C, Larsson E, Selvander H, Sjövall K, Winsa B et al. Graves' disease with moderate-severe endocrine ophthalmopathy-long term results of a prospective, randomized study of total or subtotal thyroid resection. Thyroid. 2005; 15(10):1157-1164

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

42. Jorde R, Ytre-Arne K, Størmer J, Sundsfjord J. Short-term treatment of Graves' disease with methimazole in high versus low doses. Journal of Internal Medicine. 1995; 238(2):161-165

43. Kallner G, Vitols S, Ljunggren JG. Comparison of standardized initial doses of two antithyroid drugs in the treatment of Graves' disease. Journal of Internal Medicine. 1996; 239(6):525-529

44. Kung AW, Yau CC, Cheng AC. The action of methimazole and L-thyroxine in radioiodine therapy: a prospective study on the incidence of hypothyroidism. Thyroid. 1995; 5(1):7-12

45. Leclere J. Treatment of Basedow disease with synthetic antithyroid drugs. Evaluation of the dose on the efficacy of the long term treatment. Annales d'Endocrinologie. 1994; 55(1):11-14

46. Leslie WD, Ward L, Salamon EA, Ludwig S, Rowe RC, Cowden EA. A randomized comparison of radioiodine doses in graves' hyperthyroidism. Journal of Clinical Endocrinology and Metabolism. 2003; 88(3):978-983

47. Leung AKC, Leung AAC. Evaluation and Management of Children with Thyrotoxicosis. Recent Patents on Endocrine, Metabolic & Immune Drug Discovery. 2017; 11(1):22-31

48. Li HX, Xiang N, Hu WK, Jiao XL. Relation between therapy options for Graves' disease and the course of Graves' ophthalmopathy: a systematic review and meta-analysis. Journal of Endocrinological Investigation. 2016; 39(11):1225-1233

49. Liu Y, Liu B, Liu RL, Jiang H, Huang ZN, Huang Y. A new method of subtotal thyroidectomy for Graves' disease leaving a unilateral remnant based on the upper pole. Medicine. 2017; 96(6):e5919

50. Liu ZW, Masterson L, Fish B, Jani P, Chatterjee K. Thyroid surgery for Graves' disease and Graves' ophthalmopathy. Cochrane Database of Systematic Reviews 2015, Issue 11. Art. No.: CD010576. DOI: 10.1002/14651858.CD010576.pub2.

51. Ljunggren JG, Törring O, Wallin G, Taube A, Tallstedt L, Hamberger B et al. Quality of life aspects and costs in treatment of Graves' hyperthyroidism with antithyroid drugs, surgery, or radioiodine: results from a prospective, randomized study. Thyroid. 1998; 8(8):653-659

52. Lucas A, Salinas I, Rius F, Pizarro E, Granada ML, Foz M et al. Medical therapy of Graves' disease: does thyroxine prevent recurrence of hyperthyroidism? Journal of Clinical Endocrinology and Metabolism. 1997; 82(8):2410-2413

53. Ma C, Kuang A, Xie J, Liu GJ. Radioiodine treatment for pediatric Graves' disease. Cochrane Database of Systematic Reviews 2008, Issue 3. Art. No.: CD006294. DOI: 10.1002/14651858.CD006294.pub2.

54. Ma C, Xie J, Wang H, Li J, Chen S. Radioiodine therapy versus antithyroid medications for Graves' disease. Cochrane Database of Systematic Reviews 2016, Issue 2. Art. No.: CD010094. DOI: 10.1002/14651858.CD010094.pub2.

55. Marcocci C, Bartalena L, Bogazzi F, Panicucci M, Bruno-Bossio G, Lepri A et al. Radioiodine treatment of Graves' hyperthyroidism and progression of ophthalmopathy: protective effect of systemic corticosteroids. Acta endocrinologica, supplement. 1989; 121(2):145-148

56. Mashio Y, Beniko M, Matsuda A, Koizumi S, Matsuya K, Mizumoto H et al. Treatment of hyperthyroidism with a small single daily dose of methimazole: a prospective long-term follow-up study. Endocrine Journal. 1997; 44(4):553-558

57. Mastorakos G, Doufas AG, Mantzos E, Mantzos J, Koutras DA. T4 but not T3 administration is associated with increased recurrence of Graves' disease after successful medical therapy. Journal of Endocrinological Investigation. 2003; 26(10):979-984

58. Maugendre D, Gatel A, Campion L, Massart C, Guilhem I, Lorcy Y et al. Antithyroid drugs and Graves' disease--prospective randomized assessment of long-term treatment. Clinical Endocrinology. 1999; 50(1):127-132

59. McIver B, Rae P, Beckett G, Wilkinson E, Gold A, Toft A. Lack of effect of thyroxine in patients with Graves' hyperthyroidism who are treated with an antithyroid drug. New England Journal of Medicine. 1996; 334(4):220-224

60. Menconi F, Marinò M, Pinchera A, Rocchi R, Mazzi B, Nardi M et al. Effects of total thyroid ablation versus near-total thyroidectomy alone on mild to moderate Graves' orbitopathy treated with intravenous glucocorticoids. Journal of Clinical Endocrinology and Metabolism. 2007; 92(5):1653-1658

61. Miranda-Padua ML, Cunanan EC, Kho SA, Marcelo M, Torres JF, Monzon OP et al. A randomized double-blind comparison of fixed versus calculated radioiodine dose in the treatment of graves' hyperthyroidism. Phillippine Journal of Internal Medicine. 2014; 52(3)

62. Müller PE, Bein B, Robens E, Bein HS, Spelsberg F. Thyroid surgery according to Enderlen-Hotz or Dunhill: a comparison of two surgical methods for the treatment of Graves' disease. International Surgery. 2001; 86(2):112-116

63. Nakamura H, Noh JY, Itoh K, Fukata S, Miyauchi A, Hamada N. Comparison of methimazole and propylthiouracil in patients with hyperthyroidism caused by Graves' disease. Journal of Clinical Endocrinology and Metabolism. 2007; 92(6):2157-2162

64. 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

65. Nedrebo BG, Holm PI, Uhlving S, Sorheim JI, Skeie S, Eide GE et al. Predictors of outcome and comparison of different drug regimens for the prevention of relapse in patients with Graves' disease. European Journal of Endocrinology. 2002; 147(5):583-589

66. Noh JY, Sato S, Suzuki M, Yasuda S, Matsumoto M, Kunii Y et al. Comparison of efficacy and adverse effects between methimazole 15 mg+inorganic iodine 38 mg/day and methimazole 30 mg/day as initial therapy for graves' disease patients with moderate to severe hyperthyroidism. Thyroid. 2015; 25(1):43-50

67. Orsini F, Traino AC, Grosso M, Guidoccio F, Boni G, Volterrani D et al. Personalization of radioiodine treatment for Graves' disease: a prospective, randomized study with a novel method for calculating the optimal 131I-iodide activity

based on target reduction of thyroid mass. Quarterly Journal of Nuclear Medicine and Molecular Imaging. 2012; 56(6):496-502

68. Peixoto MC, Buescu A, Goncalves MRB, Albernaz MDS, Coeli CM, Vaisman M. Antithyroid drugs for the treatment of graves disease: a randomized clinical trial. Endocrinologist. 2006; 16(6):344-348

69. Peters H, Fischer C, Bogner U, Reiners C, Schleusener H. Radioiodine therapy of Graves' hyperthyroidism: standard vs. calculated 131iodine activity. Results from a prospective, randomized, multicentre study. European Journal of Clinical Investigation. 1995; 25(3):186-193

70. Peters H, Fischer C, Bogner U, Reiners C, Schleusener H. Reduction in thyroid volume after radioiodine therapy of Graves' hyperthyroidism: results of a prospective, randomized, multicentre study. European Journal of Clinical Investigation. 1996; 26(1):59-63

71. Peters H, Fischer C, Bogner U, Reiners C, Schleusener H. Treatment of Graves' hyperthyroidism with radioiodine: results of a prospective randomized study. Thyroid. 1997; 7(2):247-251

72. Pfeilschifter J, Ziegler R. Suppression of serum thyrotropin with thyroxine in patients with Graves' disease: effects on recurrence of hyperthyroidism and thyroid volume. European Journal of Endocrinology. 1997; 136(1):81-86

73. Pirnat E, Zaletel K, Gaber??ek S, Hojker S. The outcome of 131I treatment in Graves' patients pretreated or not with methimazole. Hellenic Journal of Nuclear Medicine. 2011; 14(1):25-29

74. Pusuwan P, Tuntawiroon M, Sritongkul N, Chaudakshetrin P, Nopmaneejumruslers C, Komoltri C et al. A prospective randomized study of the efficacy and cost-effectiveness of high and low dose regimens of I-131 treatment in hyperthyroidism. Chotmaihet thangphaet [Journal of the Medical Association of Thailand]. 2011; 94(3):361-368

75. Raber W, Kmen E, Waldhäusl W, Vierhapper H. Medical therapy of Graves' disease: effect on remission rates of methimazole alone and in combination with triiodothyronine. European Journal of Endocrinology. 2000; 142(2):117-124

76. Reinwein D, Benker G, Lazarus JH, Alexander WD. A prospective randomized trial of antithyroid drug dose in Graves' disease therapy. European Multicenter Study Group on Antithyroid Drug Treatment. Journal of Clinical Endocrinology and Metabolism. 1993; 76(6):1516-1521

77. Rittmaster RS, Abbott EC, Douglas R, Givner ML, Lehmann L, Reddy S et al. Effect of methimazole, with or without L-thyroxine, on remission rates in Graves' disease. Journal of Clinical Endocrinology and Metabolism. 1998; 83(3):814-818

78. Rokni H, Sadeghi R, Moossavi Z, Treglia G, Zakavi SR. Efficacy of different protocols of radioiodine therapy for treatment of toxic nodular goiter: systematic review and meta-analysis of the literature. International Journal of Endocrinology and Metabolism. 2014; 12(2):e14424

79. Romaldini JH, Bromberg N, Werner RS, Tanaka LM, Rodrigues HF, Werner MC et al. Comparison of effects of high and low dosage regimens of antithyroid drugs in the management of Graves' hyperthyroidism. Journal of Clinical Endocrinology and Metabolism. 1983; 57(3):563-570

80. Santos RB, Romaldini JH, Ward LS. Propylthiouracil reduces the effectiveness of radioiodine treatment in hyperthyroid patients with Graves' disease. Thyroid. 2004; 14(7):525-530

81. Santos RB, Romaldini JH, Ward LS. A randomized controlled trial to evaluate the effectiveness of 2 regimens of fixed iodine (131-I) doses for Graves disease treatment. Clinical Nuclear Medicine. 2012; 37(3):241-244

82. Sapienza MT, Coura-Filho GB, Willegaignon J, Watanabe T, Duarte PS, Buchpiguel CA. Clinical and Dosimetric Variables Related to Outcome After Treatment of Graves' Disease With 550 and 1110 MBq of 131I: results of a Prospective Randomized Trial. Clinical Nuclear Medicine. 2015; 40(9):715-719

83. Schneider P, Biko J, Hänscheid H, Hilliger S, Koutsampelas C, Kranzfelder M et al. The route of administration (oral vs intravenous) does not influence dose or outcome in Graves' disease and unifocal autonomy. European Journal of Nuclear Medicine and Molecular Imaging. 2005; 32(7):788-793

84. Singhal T, Bansal S, Singhal A, McDonald S, Bal CS. Adjunctive antithyroid drugs in radioiodine therapy for hyperthyroidism. Cochrane Database of Systematic Reviews 2014, Issue 7. Art. No.: CD005447. DOI: 10.1002/14651858.CD005447.pub2.

85. Taïeb D, Bournaud C, Eberle MC, Catargi B, Schvartz C, Cavarec MB et al. Quality of life, clinical outcomes and safety of early prophylactic levothyroxine administration in patients with Graves' hyperthyroidism undergoing radioiodine therapy: a randomized controlled study. European Journal of Endocrinology. 2016; 174(4):491-502

86. Thientunyakit T, Thongmak S, Premprapha T. Comparative evaluation of two different dosage calculation protocols of iodine-131 in the treatment of hyperthyroidism. Chotmaihet thangphaet [Journal of the Medical Association of Thailand]. 2010; 93(8):969-977

87. Tian R, Kuang A. A study comparing 131I versus 131I plus antithyroid drug in the management of Graves' disease. Journal of West China University of medical sciences. 2001; 32(3):449-451

88. Unalp HR, Erbil Y, Akguner T, Kamer E, Derici H, Issever H. Does near total thyroidectomy offer advantage over total thyroidectomy in terms of postoperative hypocalcemia? International Journal of Surgery (London, England). 2009; 7(2):120-125

89. Walter MA, Christ-Crain M, Schindler C, Müller-Brand J, Müller B. Outcome of radioiodine therapy without, on or 3 days off carbimazole: a prospective interventional three-group comparison. European Journal of Nuclear Medicine and Molecular Imaging. 2006; 33(6):730-737

90. Wang J, Qin L. Radioiodine therapy versus antithyroid drugs in Graves' disease: a meta-analysis of randomized controlled trials. British Journal of Radiology. 2016; 89

91. Weetman AP, Pickerill AP, Watson P, Chatterjee VK, Edwards OM. Treatment of Graves' disease with the block-replace regimen of antithyroid drugs: the effect of treatment duration and immunogenetic susceptibility on relapse. QJM: An International journal of Medicine. 1994; 87(6):337-341

92. Witte J, Goretzki PE, Dotzenrath C, Simon D, Felis P, Neubauer M et al. Surgery for Graves' disease: total versus subtotal thyroidectomy-results of a prospective randomized trial. World Journal of Surgery. 2000; 24(11):1303-1311

93. Yousefi V, Bahreynian A, Sheikholeslami F, Tohidi M, Mehrabi Y, Azizi F. Effect of long-term continuous methimazole treatment of thyrotoxicosisi: comparison with

radioiodine. Iranian journal of endocrinology and metabolism. 2011; 12(5):466-475+555

94. Yuan J, Lu X, Yue Y. Comparison of curative effect of 131I and antithyroid drugs in Graves' disease: a meta analysis. Minerva Endocrinologica. 2018; 43(4):511-516

Appendices

Appendix A: Review protocols

Table 5:

ID	Field	Content
I	Review question	What is the clinical and cost effectiveness of using radioactive iodine vs antithyroid drugs (ATD) vs surgery to treat thyrotoxicosis secondary to Graves' disease? What is the clinical and cost effectiveness of using radioactive iodine vs surgery to treat thyrotoxicosis secondary to toxic nodular goitre? When antithyroid drugs are used, what is the most clinically and cost-effective way of using these drugs to treat thyrotoxicosis (for example choice of drugs, different treatment regimens)? When radioactive iodine is used, what is the most clinically and cost-effective way of using this treatment to treat thyrotoxicosis (for example different dosing strategies)? When surgery is indicated, what is the most clinically and cost-effective way of using surgery to treat thyrotoxicosis (for example total vs subtotal thyroidectomy)?
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	Provide clinically and cost-effective recommendations on how to manage thyrotoxicosis
IV	Eligibility criteria – population / disease / condition / issue / domain	People diagnosed with thyrotoxicosis (TSH below normal reference ranges, free T3/T4 above normal reference range)

V	Eligibility criteria – intervention(s) / exposure(s) / prognostic factor(s)	• Radioactive iodine
		  ○ Fixed administered activity strategy vs calculated absorbed radiation dose strategy
		  ○ Pre-/post- treatment with ATD vs no pre-/post- treatment
		• Antithyroid drugs
		  ○ Carbimazole/methimazole vs propylthiouracil
		  ○ Block and replace (including levothyroxine) vs titration regimen
		  ○ Duration of treatment: 6-<12 months vs 12-18 months vs >18 months
		• Surgery
		  ○ Total thyroidectomy vs subtotal thyroidectomy vs near total (Dunhill) thyroidectomy vs one sided only (hemithyroidectomy/lobectomy/isthmectomy)

VI	Eligibility criteria – comparator(s) / control or	• Comparisons between modalities • Comparisons between submodalities

	reference (gold) standard	
VII	Outcomes and prioritisation	**Critical** • Mortality (dichotomous, ≥1 year) • Quality of life (continuous) **Important (general)** • Thyroid ophthalmopathy (dichotomous) • Euthyroidism (dichotomous) • Hypothyroidism (dichotomous) • Relapse of hyperthyroidism (dichotomous) • Cardiovascular morbidity (ischaemic heart disease, dichotomous) • Arrhythmia (dichotomous) • Osteoporosis (dichotomous) • Cognitive impairment (dichotomous) • Pain (continuous) • Symptom scores (continuous) • Patient/family/carer experience (continuous) • Healthcare contacts (rates/dichotomous) **Important (surgical)** • Recurrent laryngeal nerve damage (dichotomous) • Hypocalcaemia (dichotomous) • Hypoparathyroidism (dichotomous) • Bleeding (dichotomous) • Infection (dichotomous) **Important (pharmacological)** • Agranulocytosis (dichotomous) • Liver failure (dichotomous) • Minor drug related adverse effects (dichotomous) • Teratogenesis (dichotomous) **Important (radioiodine)** • Infertility (dichotomous) • Malignancy (dichotomous) • Thyrotoxic storm (dichotomous) • Growth abnormalities (dichotomous) • Hypocalcaemia (dichotomous) • Hypoparathyroidism (dichotomous) • Teratogenesis (dichotomous) Minimum duration as for the minimum duration for inclusion of studies unless specified.
VIII	Eligibility criteria – study design	• Minimum follow-up of 3 months • RCTs • Non-randomised cohort studies to be considered if adjusted for key confounders (age, co-existing conditions, baseline T4, size of goitre) and insufficient RCTs evidence found, on an intervention by intervention basis
IX	Other inclusion / exclusion criteria	• Excluding studies in pregnancy • Excluding studies aimed specifically at treating thyroid eye disease • Excluding studies in context of thyroid malignancy
X	Proposed	**Stratifications**

	sensitivity / subgroup analysis, or meta-regression	• Age – young children (0-4), children and young people (4-18), adults (>18-65), older adults (>65) • For antithyroid drugs vs radioactive iodine vs surgery - Cause of thyrotoxicosis (Graves' disease, toxic nodular goitre, thyroiditis) • Treatment stage – naïve/general (non-naïve, downgraded for indirectness), second line (remain symptomatic despite previous treatment, as defined by studies) **Subgroup analyses** • Gender (male only vs female only) • Age subdivisions (4-12, 12-18, 18-50, 50-65, 65-85, >85) • Comparison not under investigation (for example for block and replace vs titration, if some studies use methimazole and others use propylthiouracil)
XI	Selection process – duplicate screening / selection / analysis	• A sample of at least 10% of the abstract lists were double-sifted by a senior research fellow and discrepancies rectified, with committee input where consensus could not be reached, 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 / duplicate	A standardised evidence table format will be used and published as an appendix of the evidence report.
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 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

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 – 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).[64] **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.

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.

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.

Table 7: Database date parameters and filters used

Database	Dates searched	Search filter used
Medline (OVID)	1946 – 07 January 2019	Exclusions Randomised controlled trials Systematic review studies Observational studies
Embase (OVID)	1974 – 07 January 2019	Exclusions Randomised controlled trials Systematic review studies Observational 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 goiter/
2.	exp Hyperthyroidism/
3.	(hyperthyroid* or thyrotoxicosis).ti,ab.
4.	(toxic adj4 (node* or nodul* or multi?nodul* or goitre or goiter)).ti,ab.
5.	(graves' disease or plummer's disease).ti,ab.
6.	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.	randomized controlled trial.pt.
26.	controlled clinical trial.pt.
27.	randomi#ed.ti,ab.
28.	placebo.ab.
29.	randomly.ti,ab.
30.	Clinical Trials as topic.sh.
31.	trial.ti.
32.	or/25-31
33.	Meta-Analysis/
34.	exp Meta-Analysis as Topic/
35.	(meta analy* or metanaly* or metaanaly* or meta regression).ti,ab.
36.	((systematic* or evidence*) adj3 (review* or overview*)).ti,ab.
37.	(reference list* or bibliograph* or hand search* or manual search* or relevant journals).ab.
38.	(search strategy or search criteria or systematic search or study selection or data extraction).ab.
39.	(search* adj4 literature).ab.
40.	(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.
41.	cochrane.jw.
42.	((multiple treatment* or indirect or mixed) adj2 comparison*).ti,ab.
43.	or/33-42
44.	Epidemiologic studies/
45.	Observational study/
46.	exp Cohort studies/
47.	(cohort adj (study or studies or analys* or data)).ti,ab.
48.	((follow up or observational or uncontrolled or non randomi#ed or epidemiologic*) adj (study or studies or data)).ti,ab.
49.	((longitudinal or retrospective or prospective or cross sectional) and (study or studies or review or analys* or cohort* or data)).ti,ab.
50.	Controlled Before-After Studies/
51.	Historically Controlled Study/
52.	Interrupted Time Series Analysis/
53.	(before adj2 after adj2 (study or studies or data)).ti,ab.
54.	or/4-53
55.	exp case control study/
56.	case control*.ti,ab.

57.	or/55-56
58.	54 or 57
59.	Cross-sectional studies/
60.	(cross sectional and (study or studies or review or analys* or cohort* or data)).ti,ab.
61.	or/59-60
62.	54 or 61
63.	54 or 57 or 61
64.	6 not 24
65.	limit 64 to English language
66.	65 and (32 or 43 or 64)

Embase (Ovid) search terms

1.	goiter/
2.	hyperthyroidism/ or graves disease/ or thyrotoxicosis/ or toxic goiter/
3.	(hyperthyroid* or thyrotoxicosis).ti,ab.
4.	(toxic adj4 (node* of nodul* or multi?nodul* or goitre or goiter)).ti,ab.
5.	(graves' disease or plummer's disease).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.	random*.ti,ab.
25.	factorial*.ti,ab.
26.	(crossover* or cross over*).ti,ab.
27.	((doubl* or singl*) adj blind*).ti,ab.
28.	(assign* or allocat* or volunteer* or placebo*).ti,ab.
29.	crossover procedure/
30.	single blind procedure/
31.	randomized controlled trial/
32.	double blind procedure/
33.	or/24-32
34.	systematic review/

35.	meta-analysis/
36.	(meta analy* or metanaly* or metaanaly* or meta regression).ti,ab.
37.	((systematic* or evidence*) adj3 (review* or overview*)).ti,ab.
38.	(reference list* or bibliograph* or hand search* or manual search* or relevant journals).ab.
39.	(search strategy or search criteria or systematic search or study selection or data extraction).ab.
40.	(search* adj4 literature).ab.
41.	(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.
42.	cochrane.jw.
43.	((multiple treatment* or indirect or mixed) adj2 comparison*).ti,ab.
44.	or/34-43
45.	Clinical study/
46.	Observational study/
47.	family study/
48.	longitudinal study/
49.	retrospective study/
50.	prospective study/
51.	cohort analysis/
52.	follow-up/
53.	cohort*.ti,ab.
54.	52 and 53
55.	(cohort adj (study or studies or analys* or data)).ti,ab.
56.	((follow up or observational or uncontrolled or non randomi#ed or epidemiologic*) adj (study or studies or data)).ti,ab.
57.	((longitudinal or retrospective or prospective or cross sectional) and (study or studies or review or analys* or cohort* or data)).ti,ab.
58.	(before adj2 after adj2 (study or studies or data)).ti,ab.
59.	or/45-51,54-58
60.	exp case control study/
61.	case control*.ti,ab.
62.	or/60-61
63.	59 or 62
64.	cross-sectional study/
65.	(cross sectional and (study or studies or review or analys* or cohort* or data)).ti,ab.
66.	or/64-65
67.	59 or 66
68.	59 or 62 or 66
69.	23 and (33 or 44 or 68)
70.	limit 69 to English language

Cochrane Library (Wiley) search terms

#1.	MeSH descriptor: [Goiter] explode all trees
#2.	MeSH descriptor: [Hyperthyroidism] explode all trees
#3.	(hyperthyroid* or thyrotoxicosis):ti,ab
#4.	(toxic near/4 (node* or nodul* or multinodul* or multi-nodul* or goitre or goiter)):ti,ab

#5.	MeSH descriptor: [Graves Disease] explode all trees
#6.	(grave* near/4 (thyrotoxicos* or hyperthyr*)):ti,ab
#7.	graves' disease:ti,ab
#8.	(or #1-#7)

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 8: 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.

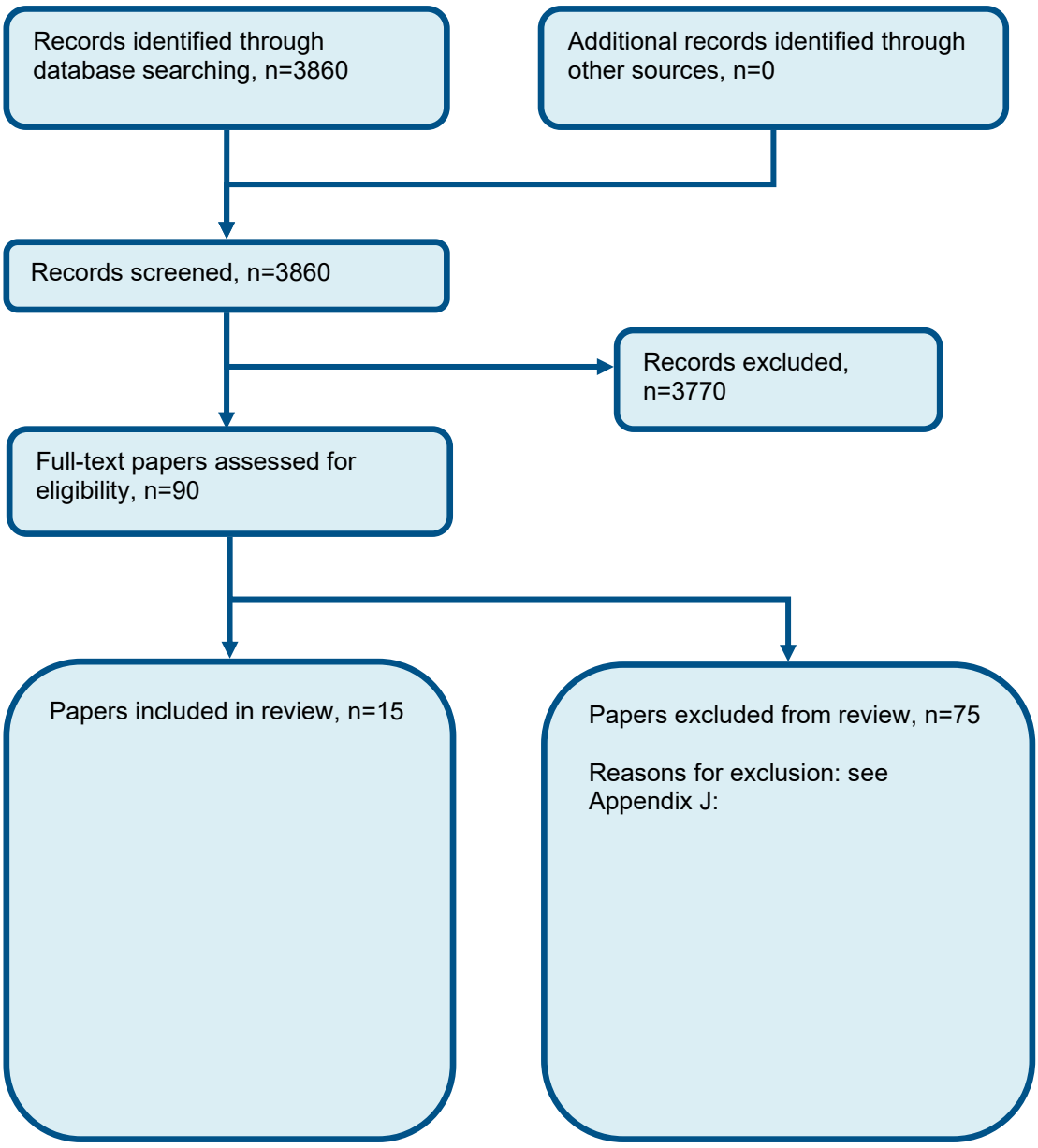
67.	(willingness to pay or time tradeoff or time trade off or tto or standard gamble*).ti,ab.
68.	(sf36* or sf 36* or short form 36* or shortform 36* or shortform36*).ti,ab.
69.	(sf20 or sf 20 or short form 20 or shortform 20 or shortform20).ti,ab.
70.	(sf12* or sf 12* or short form 12* or shortform 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)

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

Appendix C: Clinical evidence selection

Figure 1: Flow chart of clinical study selection for the review of thyrotoxicosis (drugs)



Appendix D: Clinical evidence tables

Study	Allannic 1990[3]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=114)
Countries and setting	Conducted in France; Setting:
Line of therapy	1st line
Duration of study	Follow up (post intervention): 24 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: based on usual clinical signs and symptoms including ophthalmopathy and hyperthyroidism, measurement of serum thyroid hormone concentrations
Stratum	Graves' disease
Subgroup analysis within study	Not applicable
Inclusion criteria	patients with hyperthyroidism due to Graves' disease, examined for the first time, with no previous treatment for this affection
Exclusion criteria	pregnancy, toxic nodular goiters
Recruitment/selection of patients	All patients at institution with hyperthyroidism due to Graves' disease
Age, gender and ethnicity	Age - Mean (SD): 18-month treatment: 39.2 (12.3); 6-month treatment: 43.1 (14.7). Gender (M:F): 15/79. Ethnicity: not specified
Further population details	1. Age: 2. Gender:
Indirectness of population	No indirectness
Interventions	(n=57) Intervention 1: 12-18 month treatment. 30-60 mg/d carbimazole, later reduced to 10-20 mg/d to maintain euthyroidism. Duration 18 months. Concurrent medication/care: not specified. Indirectness: No indirectness (n=57) Intervention 2: 6-<12 month treatment. 30-60 mg/d carbimazole, later reduced to 10-20 mg/d to maintain euthyroidism. Duration 6 months. Concurrent medication/care: not specified. Indirectness: No indirectness

Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: 12-18 MONTH TREATMENT versus 6-<12 MONTH TREATMENT Protocol outcome 1: Euthyroidism - Actual outcome for Graves' disease: Remission (clinical euthyroidism) at 24 months after treatment; Group 1: 29/46, Group 2: 20/48 Risk of bias: All domain - Low, 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: 11, Reason: 7 were excluded for not complying with the treatment and follow-up protocols, 4 were excluded since they were not rechecked during the 2 year follow-up period; Group 2 Number missing: 9, Reason: 7 were excluded for not complying with the treatment and follow-up protocols, 2 were excluded since they were not rechecked during the 2 year follow-up period Protocol outcome 2: Relapse of hyperthyroidism - Actual outcome for Graves' disease: Relapse at 24 months after treatment; Group 1: 17/46, Group 2: 28/48 Risk of bias: All domain - Low, 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: 11, Reason: 7 were excluded for not complying with the treatment and follow-up protocols, 4 were excluded since they were not rechecked during the 2 year follow-up period; Group 2 Number missing: 9, Reason: 7 were excluded for not complying with the treatment and follow-up protocols, 2 were excluded since they were not rechecked during the 2 year follow-up period	
Protocol outcomes not reported by the study	Quality of life; Mortality; Thyroid ophthalmopathy; Hypothyroidism; Ischaemic heart disease; Heart failure; Arrhythmia; Osteoporosis; Impaired cognitive function; Growth; Pain; Symptom scores; Experience of care; Healthcare contacts; Agranulocytosis; Liver failure; Minor drug related adverse events; Teratogenesis

Study	**Edmonds 1994[27]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=70)
Countries and setting	Conducted in United Kingdom; Setting: endocrine clinic
Line of therapy	1st line
Duration of study	Intervention + follow up: 12-month treatment, 24-month follow up
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: thyroid function tests
Stratum	Graves' disease
Subgroup analysis within study	Not applicable
Inclusion criteria	patients with newly diagnosed untreated Graves' disease, age range 15-65, who were not pregnant and in whom medical therapy was indicated. Diagnosis based on measurements of plasma free T3 and TSH and the demonstration of a diffusely increased thyroid uptake of 99mTcO4 (pertechnetate)
Exclusion criteria	not specified
Recruitment/selection of patients	referral to clinic
Age, gender and ethnicity	Age - Mean (SD): block-replace: 48(11.9); titration: 41(12.9). Gender (M:F): 22/95. Ethnicity: European, Asian
Further population details	1. Age: 2. Gender:
Indirectness of population	No indirectness
Interventions	(n=49) Intervention 1: Block and replace. carbimazole 60 mg/d, T4 100-150 µg/d (beginning at 4 weeks after carbimazole). Duration 12 months. Concurrent medication/care: not specified. Indirectness: No indirectness (n=46) Intervention 2: Titration. carbimazole 60 mg/d for four weeks, then reduced to reach maintenance dose (usually by the third month of treatment). Duration 12 months. Concurrent medication/care: not specified. Indirectness: No indirectness
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: BLOCK AND REPLACE versus TITRATION

Protocol outcome 1: Relapse of hyperthyroidism

- Actual outcome for Graves' disease: Relapse of hyperthyroidism at 24 months after treatment completion; Group 1: 17/34, Group 2: 24/36 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 - Low; Indirectness of outcome: No indirectness ; Baseline details: small, reportedly not significant differences in age, race, gender; Group 1 Number missing: 15, Reason: failed to complete treatment: changed to propylthiouracil (n=6), relapsed (n=3), were treated with radioiodine (n=3), had partial thyroidectomy (n=3); Group 2 Number missing: 10, Reason: failed to complete treatment: changed to propylthiouracil (n=6), relapsed (n=3), were treated with radioiodine (n=1), had partial thyroidectomy (n=3) Protocol outcome 2: Agranulocytosis - Actual outcome for Graves' disease: Agranulocytosis at 3 weeks of treatment; Group 1: 1/49, Group 2: 0/46 Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: small, reportedly not significant differences in age, race, gender; Group 1 Number missing: ; Group 2 Number missing: Protocol outcome 3: Minor drug related adverse events - Actual outcome for Graves' disease: side effects leading to withdrawal at during treatment; Group 1: 7/49, Group 2: 6/46 Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: small, reportedly not significant differences in age, race, gender; Group 1 Number missing: ; Group 2 Number missing:	
Protocol outcomes not reported by the study	Quality of life; Mortality; Thyroid ophthalmopathy; Euthyroidism; Hypothyroidism; Ischaemic heart disease; Heart failure; Arrhythmia; Osteoporosis; Impaired cognitive function; Growth; Pain; Symptom scores; Experience of care; Healthcare contacts; Liver failure; Teratogenesis

Study	**García-mayor 1992[29]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=52)
Countries and setting	Conducted in Unknown; Setting: outpatients
Line of therapy	1st line
Duration of study	Intervention time: 12 or 24 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: laboratory tests
Stratum	Graves' disease
Subgroup analysis within study	Not applicable: Overall
Inclusion criteria	Symptoms and signs of hyperthyroidism, elevated T4, FT4 and suppressed level of sTSH, diffuse up-take of technetium 99, antibodies to TSH receptor (TRAb) over normal value (<15 U/L, defined as mean ±2 SD of data obtained with sera of 75 healthy controls)
Exclusion criteria	Pregnancy, enlarged goiter which required surgery because of local problems and patients who required L-T4 administration in order to prevent hypothyroidism within the period of treatment.
Recruitment/selection of patients	Thyroid Unit attendees
Age, gender and ethnicity	Age - Mean (SD): 39.35 (13.69). Gender (M:F): 3/49. Ethnicity: Not specified
Further population details	1. Age: 2. Gender:
Extra comments	Graves' disease
Indirectness of population	No indirectness
Interventions	(n=29) Intervention 1: 12-18 month treatment. 10 mg carbimazole every 8 h., reduced to no less than 10 mg/day once euthyroid state reached. Duration 12 months. Concurrent medication/care: not specified. Indirectness: No indirectness (n=27) Intervention 2: >18-month treatment. 10 mg carbimazole every 8 h., reduced to no less than 10 mg/day once euthyroid state reached. Duration 24 months. Concurrent medication/care: not specified. Indirectness: No indirectness
Funding	Academic or government funding (Spanish Ministry of Health)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: 12-18 MONTH TREATMENT versus >18 MONTH TREATMENT	

Protocol outcome 1: Relapse of hyperthyroidism - Actual outcome for Graves' disease: Relapse (elevated FT4 and suppressed sTSH levels with or without elevated TRAb levels) at 5 years after stopping drug therapy; Group 1: 13/28, Group 2: 13/24 Risk of bias: All domain - High, Selection - Very high, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: baseline comparability for thyroid hormone levels not specified; Group 1 Number missing: 1, Reason: not completing treatment and/or follow-up protocol or not being rechecked during the 5 year follow-up; Group 2 Number missing: 3, Reason: not completing treatment and/or follow-up protocol or not being rechecked during the 5 year follow-up	
Protocol outcomes not reported by the study	Quality of life; Mortality; Thyroid ophthalmopathy; Euthyroidism; Hypothyroidism; Ischaemic heart disease; Heart failure; Arrhythmia; Osteoporosis; Impaired cognitive function; Growth; Pain; Symptom scores; Experience of care; Healthcare contacts; Agranulocytosis; Liver failure; Minor drug related adverse events; Teratogenesis

Study	Grebe 1998[32]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=37)
Countries and setting	Conducted in New Zealand; Setting: outpatients
Line of therapy	1st line
Duration of study	Intervention + follow up: 6 -month intervention + 24-month follow-up
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: clinical assessment
Stratum	Graves' disease
Subgroup analysis within study	Not applicable
Inclusion criteria	Patients presenting with first episode of Graves' disease to the Wellington Hospital department of Endocrinology. Diagnosis of Graves' disease defined as clinical and biochemical evidence of thyrotoxicosis associated with a smooth goitre with uniformally increased 99mTc uptake.
Exclusion criteria	known pituitary, liver or haematological abnormalities; pregnancy, known allergies to thionamide drugs
Recruitment/selection of patients	Consecutive
Age, gender and ethnicity	Age - Mean (SD): Block-replace: 33 (8.7); Titration: 33.7 (11.9). Gender (M:F): 8/29. Ethnicity: European (n=26), Polynesian (n=4), Chinese (n=6), Middle-Eastern (n=1)
Further population details	1. Age: 2. Gender:
Indirectness of population	No indirectness
Interventions	(n=17) Intervention 1: Block and replace. carbimazole 100 mg/d, T4 starting 2-3 weeks later adjusted to maintain euthyroidism (serum thyroid function test results within reference range), mean T4 dose increased to 132 µg/d at end of treatment. Duration 6 months. Concurrent medication/care: not specified. Indirectness: No indirectness (n=20) Intervention 2: Titration. carbimazole 25 mg/d titrated to maintain euthyroidism (average 17 mg/d at end of treatment). Duration 6 months. Concurrent medication/care: not specified. Indirectness: No indirectness
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: BLOCK AND REPLACE versus TITRATION	

Protocol outcome 1: Relapse of hyperthyroidism - Actual outcome for Graves' disease: Relapse at 24 months after treatment withdrawal; Group 1: 13/16, Group 2: 16/17 Risk of bias: All domain - High, Selection - Very high, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: differences in goitre size, FT4 levels and total white blood cell count; Group 1 Number missing: 1, Reason: lost to follow up; Group 2 Number missing: 3, Reason: lost to follow up (n=1), elected to continue with carbimazole treatment and were excluded from analysis (n=2) Protocol outcome 2: Agranulocytosis - Actual outcome for Graves' disease: Agranulocytosis at during the 6 month treatment; Group 1: 2/17, Group 2: 1/20 Risk of bias: All domain - High, Selection - Very high, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: differences in goitre size, FT4 levels and total white blood cell count; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 3: Minor drug related adverse events - Actual outcome for Graves' disease: Treatment side-effects (skin reactions and other side-effects) at during the 6 month treatment; Group 1: 5/17, Group 2: 0/20 Risk of bias: All domain - Very high, Selection - Very high, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: differences in goitre size, FT4 levels and total white blood cell count; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Quality of life; Mortality; Thyroid ophthalmopathy; Euthyroidism; Hypothyroidism; Ischaemic heart disease; Heart failure; Arrhythmia; Osteoporosis; Impaired cognitive function; Growth; Pain; Symptom scores; Experience of care; Healthcare contacts; Liver failure; Teratogenesis

Study	He 2004[35]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=30)
Countries and setting	Conducted in China; Setting: outpatients
Line of therapy	1st line
Duration of study	Intervention time: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: according to history and signs of hyperthyroidism
Stratum	Graves' disease
Subgroup analysis within study	Not applicable
Inclusion criteria	newly diagnosed Graves' hyperthyroidism
Exclusion criteria	not specified
Recruitment/selection of patients	Randomly
Age, gender and ethnicity	Age - Mean (SD): MMI: 32 (7.1); PTU: 31(6.5). Gender (M:F): 9/21. Ethnicity: Not stated
Further population details	1. Age: 2. Gender:
Indirectness of population	No indirectness
Interventions	(n=15) Intervention 1: Carbimazole/methimazole. 15 mg/d. Duration 12 weeks. Concurrent medication/care: not specified. Indirectness: No indirectness (n=15) Intervention 2: Propylthiouracil. 150 mg/d. Duration 12 weeks. Concurrent medication/care: Not specified. Indirectness: No indirectness
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CARBIMAZOLE/METHIMAZOLE versus PROPYLTHIOURACIL

Protocol outcome 1: Euthyroidism
- Actual outcome for Graves' disease: Euthyroidism at 12 weeks; Group 1: 12/15, Group 2: 5/15
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: 0; Group 2 Number missing: 0 Protocol outcome 2: Hypothyroidism - Actual outcome for Graves' disease: hypothyroidism at 12 weeks; Group 1: 4/15, Group 2: 0/15 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: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Quality of life; Mortality; Thyroid ophthalmopathy; Relapse of hyperthyroidism; Ischaemic heart disease; Heart failure; Arrhythmia; Osteoporosis; Impaired cognitive function; Growth; Pain; Symptom scores; Experience of care; Healthcare contacts; Agranulocytosis; Liver failure; Minor drug related adverse events; Teratogenesis

Study	Homsanit 2001[37]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=71)
Countries and setting	Conducted in Thailand; Setting: not specified
Line of therapy	1st line
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Clinical assessment and specified criteria
Stratum	Graves' disease
Subgroup analysis within study	Not applicable
Inclusion criteria	newly diagnosed Graves' hyperthyroidism, based on clinical and biochemical features including diffuse enlargement of thyroid gland, presence of signs and symptoms of thyrotoxicosis, and elevated serum thyroid hormones accompanied with suppressed serum TSH levels
Exclusion criteria	other common causes of thyrotoxicosis i.e. toxic multinodular goitre, toxic adenoma and thyroiditis, concomitant medications known to interfere with thyroid hormone metabolism including thyroxin, β-adrenergic blocking agents, lithium, amiodarone, glucocorticoids, oral contraceptive pills and other oestrogen containing agents.
Recruitment/selection of patients	not specified
Age, gender and ethnicity	Age - Mean (SD): MMI: 35.4 (11.5); PTU: 34.8 (13). Gender (M:F): 9/62. Ethnicity: Not specified
Further population details	1. Age: 2. Gender:
Indirectness of population	No indirectness
Interventions	(n=35) Intervention 1: Carbimazole/methimazole. 15 mg/d. Duration 12 weeks. Concurrent medication/care: no concomitant treatment. Indirectness: No indirectness (n=36) Intervention 2: Propylthiouracil. 150 mg/d. Duration 12 weeks. Concurrent medication/care: no concomitant treatment. Indirectness: No indirectness
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CARBIMAZOLE/METHIMAZOLE versus PROPYLTHIOURACIL	

Protocol outcome 1: Euthyroidism - Actual outcome for Graves' disease: Euthyroidism at 12 weeks; Group 1: 27/35, Group 2: 7/36 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: 0; Group 2 Number missing: 0 Protocol outcome 2: Hypothyroidism - Actual outcome for Graves' disease: Hypothyroidism at 12 weeks; Group 1: 11/35, Group 2: 0/36 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: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Quality of life; Mortality; Thyroid ophthalmopathy; Relapse of hyperthyroidism; Ischaemic heart disease; Heart failure; Arrhythmia; Osteoporosis; Impaired cognitive function; Growth; Pain; Symptom scores; Experience of care; Healthcare contacts; Agranulocytosis; Liver failure; Minor drug related adverse events; Teratogenesis

Study	Lucas 1997[52]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=60)
Countries and setting	Conducted in Spain; Setting: outpatients
Line of therapy	1st line
Duration of study	Intervention + follow up: 12-24 month intervention, mean (SD) intervention time 18.4 (2.6) months, mean (SD) follow up time 4.98(1.6) years.
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: clinical examination including measurement of serum total T3, T4, free T4 and TSH
Stratum	Graves' disease
Subgroup analysis within study	Not applicable
Inclusion criteria	untreated patients with initial episode of Graves' disease (GD)hyperthyroidism, living in area of normal iodine intake. Diagnosis of GD based on measurements of serum total T3, T4, free T4 and TSH, nodulation absence by thyroid palpation and demonstration of a diffuse increased thyroid uptake of 99m TcO4 (pertechnetate)
Exclusion criteria	not specified
Recruitment/selection of patients	consecutive
Age, gender and ethnicity	Age - Mean (SD): Block replace: 34.5(8.3); Titration: 37.5(13.9). Gender (M:F): 11/49. Ethnicity: Spanish
Further population details	1. Age: 2. Gender:
Indirectness of population	No indirectness
Interventions	(n=30) Intervention 1: Block and replace. carbimazole 30-45 mg/d, T4 100µg/d, adjusted after 1 month to 75-150 µg/d to maintain euthyroidism. Duration 12-24 months mean (SD): 18.4(2.6). Concurrent medication/care: carbimazole 45-60 mg/d until euthyroid. Indirectness: No indirectness (n=30) Intervention 2: Titration. initial carbimazole dose 45-60 mg/d adjusted to maintain euthyroidism. Duration 12-24 months mean (SD): 18.4(2.6). Concurrent medication/care: carbimazole 45-60 mg/d until euthyroid. Indirectness: No indirectness
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: BLOCK AND REPLACE versus TITRATION Protocol outcome 1: Relapse of hyperthyroidism - Actual outcome for Graves' disease: Relapse at mean (SD) 8.5 (8.7) months after carbimazole withdrawal; 4.98 (1.6) year follow up; Group 1: 20/30, Group 2: 18/30 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 ; Baseline details: difference in frequency of evolution of symptomatology; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Quality of life; Mortality; Thyroid ophthalmopathy; Euthyroidism; Hypothyroidism; Ischaemic heart disease; Heart failure; Arrhythmia; Osteoporosis; Impaired cognitive function; Growth; Pain; Symptom scores; Experience of care; Healthcare contacts ; Agranulocytosis ; Liver failure ; Minor drug related adverse events ; Teratogenesis

Study	**Maugendre 1999[58]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=175)
Countries and setting	Conducted in France; Setting: not specified
Line of therapy	1st line
Duration of study	Follow up (post intervention): 24 months
Method of assessment of guideline condition	Inadequate method of assessment/diagnosis: diagnosis was based on 'the usual clinical signs', laboratory tests were only performed in patients who consulted for their first episode of hyperthyroidism
Stratum	Graves' disease
Subgroup analysis within study	Not applicable
Inclusion criteria	patients with Graves' disease, diagnosis based on the usual clinical signs including hyperthyroidism and ophthalmopathy, and laboratory tests in patients consulted for their first episode of hyperthyroidism
Exclusion criteria	patients with toxic nodular goitres, previous treatment of Graves' disease
Recruitment/selection of patients	not specified
Age, gender and ethnicity	Age - Median (range): 40.4 (13-74). Gender (M:F): 19/115. Ethnicity: French
Further population details	1. Age: 2. Gender:
Extra comments	Caucasian
Indirectness of population	No indirectness
Interventions	(n=82) Intervention 1: >18-month treatment. 20-50 mg/d carbimazole for 3 months, followed by 10-15 mg/d to maintain euthyroidism . Duration 42 months. Concurrent medication/care: not specified. Indirectness: No indirectness (n=93) Intervention 2: 12 - 18-month treatment. 20-50 mg/d carbimazole for 3 months, followed by 10-15 mg/d to maintain euthyroidism . Duration 18 months. Concurrent medication/care: not specified. Indirectness: No indirectness
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: >18 MONTH TREATMENT versus 12-18 MONTH TREATMENT	

Protocol outcome 1: Relapse of hyperthyroidism - Actual outcome for Graves' disease: Relapse at 24 months; Group 1: 18/62, Group 2: 26/72 Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Groups differed in Age; Group 1 Number missing: 20, Reason: 2 patients preferred radical therapy shortly after randomization, 1 did not respond to drug therapy, 1 developed adverse events, 8 did not comply with the treatment protocol, 8 were lost to follow-up; Group 2 Number missing: 21, Reason: 1 patients preferred radical therapy shortly after randomization, 1 did not respond to drug therapy, 4 developed adverse events, 6 did not comply with the treatment protocol, 9 were lost to follow-up	
Protocol outcomes not reported by the study	Quality of life; Mortality; Thyroid ophthalmopathy; Euthyroidism; Hypothyroidism; Ischaemic heart disease; Heart failure; Arrhythmia; Osteoporosis; Impaired cognitive function; Growth; Pain; Symptom scores; Experience of care; Healthcare contacts; Agranulocytosis; Liver failure; Minor drug related adverse events; Teratogenesis

Study	Mciver 1996[59]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=111)
Countries and setting	Conducted in United Kingdom; Setting:
Line of therapy	1st line
Duration of study	Follow up (post intervention): median 12 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: diagnosis based clinical examination involving blood test
Stratum	Graves' disease
Subgroup analysis within study	Not applicable
Inclusion criteria	previously untreated patients with hyperthyroidism due to Graves' disease, diagnosed on the basis of elevated concentrations of free T4, total T3 and undetectable TSH (<0.04 µU/mL), presence of diffuse goitre, ophthalmopathy or pretibial myxedema, or detectable serum concentrations of thyrotropin-receptor antibodies
Exclusion criteria	not specified
Recruitment/selection of patients	consecutive
Age, gender and ethnicity	Age - Mean (SD): block-replace: 36(10); titration: 33(9). Gender (M:F): 22/89. Ethnicity:
Further population details	1. Age: 2. Gender:
Indirectness of population	No indirectness
Interventions	(n=59) Intervention 1: Block and replace. carbimazole 20 mg twice daily; T4 initially 100 mg/d adjusted to achieve undetectable TSH (<0.04 µU/ml). Duration 17 months combination; 18 months T4 alone. Concurrent medication/care: carbimazole 40mg/d for 1 month. Indirectness: No indirectness (n=52) Intervention 2: Titration. carbimazole started at 40mg/d adjusted to achieve normal TSH, T4 and T3. Duration 17 months (total 18 months). Concurrent medication/care: carbimazole 40mg/d for 1 month. Indirectness: No indirectness
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: BLOCK AND REPLACE versus TITRATION

Protocol outcome 1: Relapse of hyperthyroidism

- Actual outcome for Graves' disease: Recurrence of hyperthyroidism at 3-18 months post treatment withdrawal (12 month median follow-up); Group 1: 8/25, Group 2: 8/20; Comments: 10 patients in each group were withdrawn from the study. Follow-up data was only available in 53 participants. Number of people analysed within each group is not given. Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Very high, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - High; Indirectness of outcome: No indirectness ; Group 1 Number missing: 34, Reason: withdrawal from study due to side effects of drug (urticaria, arthralgia or nausea, noncompliance, change in residence, loss at follow up; Group 2 Number missing: 32, Reason: withdrawal from study due to side effects of drug (urticaria, arthralgia or nausea, noncompliance, change in residence, loss at follow up	
Protocol outcomes not reported by the study	Quality of life; Mortality; Thyroid ophthalmopathy; Euthyroidism; Hypothyroidism; Ischaemic heart disease; Heart failure; Arrhythmia; Osteoporosis; Impaired cognitive function; Growth; Pain; Symptom scores; Experience of care; Healthcare contacts; Agranulocytosis; Liver failure; Minor drug related adverse events; Teratogenesis

Study	**Nakamura 2007[63]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=396)
Countries and setting	Conducted in Japan; Setting: outpatients
Line of therapy	1st line
Duration of study	Intervention time: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: diagnosed according to Japan thyroid association's diagnosis guidelines
Stratum	Graves' disease
Subgroup analysis within study	Not applicable
Inclusion criteria	patients with untreated hyperthyroidism due to GD, diagnosed according to Japan Thyroid Association's diagnosis guidelines
Exclusion criteria	Age younger than 16-year old; pregnancy; relapsed patients after subtotal thyroidectomy or radioiodine therapy; previous treatment with ATD; severe complications such as heart failure; and patients on glucocorticoid steroids or drugs that may influence thyroid functions
Recruitment/selection of patients	patients seen by four different hospitals
Age, gender and ethnicity	Age - Mean (SD): MMI: 40.29 (13.3); PTU: 40.2 (12.9). Gender (M:F): 63/240. Ethnicity: Not specified
Further population details	1. Age: 2. Gender:
Indirectness of population	No indirectness
Interventions	(n=282) Intervention 1: Carbimazole/methimazole. 15 (single dose) to 30 (two divided doses) mg/d; lessened to 10 or 15 mg/d when normal FT4 (0.8 - 1.6 ng/dl) and FT3 (3.1-4.9 pg/ml) at weeks 4 and 8. Duration 12 weeks. Concurrent medication/care: β- blocker given when necessary. Indirectness: No indirectness (n=114) Intervention 2: Propylthiouracil. 300 mg/d (three divided doses), lessened to 150 mg when normal FT4 (0.8 - 1.6 ng/dl) and FT3 (3.1-4.9 pg/ml) at weeks 4 and 8. Duration 12 weeks. Concurrent medication/care: β-blocker given when necessary. Indirectness: No indirectness
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CARBIMAZOLE/METHIMAZOLE versus PROPYLTHIOURACIL	

Protocol outcome 1: Euthyroidism - Actual outcome for Graves' disease: Euthyroidism based on: normal FT4/FT3 at 12 weeks; Group 1: 176/194, Group 2: 54/69 Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 88, Reason: drop out, side effects, not visiting regularly; Group 2 Number missing: 45, Reason: drop out, side effects, not visiting regularly Protocol outcome 2: Minor drug related adverse events - Actual outcome for Graves' disease: Drug related adverse effects at 12 weeks; Group 1: 58/267, Group 2: 54/104; Comments: Hepatotoxicity, skin eruption/urticaria, leukocytopenia or other Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness; Group 1 Number missing: 15, Reason: drop out patients excluded; Group 2 Number missing: 10, Reason: drop out patients excluded	
Protocol outcomes not reported by the study	Quality of life; Mortality; Thyroid ophthalmopathy; Hypothyroidism; Relapse of hyperthyroidism; Ischaemic heart disease; Heart failure; Arrhythmia; Osteoporosis; Impaired cognitive function; Growth; Pain; Symptom scores; Experience of care; Healthcare contacts; Agranulocytosis; Liver failure; Teratogenesis

Study	Nedrebo 2002[65]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=218)
Countries and setting	Conducted in Norway; Setting: not specified
Line of therapy	1st line
Duration of study	Intervention + follow up: 12-months + 24-month follow-up
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: clinical assessment
Stratum	Graves' disease
Subgroup analysis within study	Not applicable
Inclusion criteria	patients with Graves' disease between 16 and 75 years of age. Diagnosis based on the clinical signs of hyperthyroidism combined with suppressed serum TSH and positive TRAb or ophthalmopathy. Recruited from four hospitals in Norway
Exclusion criteria	pregnancy, treatment with antithyroid drugs (ATD) in the 12 months prior to enrollment, allergy to ATD, ongoing immunosuppressive treatment, non-compliance because of psychiatric or other serious diseases, patients' preference for surgery or radioiodine treatment, or unwillingness to participate in the study
Recruitment/selection of patients	consecutive
Age, gender and ethnicity	Age - Mean (SD): block-replace: 42.02 (11.04); titration: 42.8 (12.77). Gender (M:F): 30/188. Ethnicity: Caucasian (n=214), Asiatic (n=4)
Further population details	1. Age: 2. Gender:
Indirectness of population	No indirectness
Interventions	(n=110) Intervention 1: Block and replace. carbimazole at initial mean dose (range) 29.7 mg/d (15-45 mg) except for one patient receiving propylthiouracil 200-400 mg/d; L-T4 to maintain normal FT-4 once euthyroid. Duration 12 months. Concurrent medication/care: Beta-blockers given initially according to clinical judgment. Indirectness: No indirectness (n=108) Intervention 2: Titration. carbimazole at initial mean dose (range) 29.7 mg/d (15-45 mg) except for five patients receiving propylthiouracil 200-400 mg/d; initial dose adjusted to maintain normal serum FT4 once euthyroid. Duration 12 months. Concurrent medication/care: Beta-blockers given initially according to clinical judgment. Indirectness: No indirectness

Funding	Academic or government funding (Norwegian Research Council, Helse Vest)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: BLOCK AND REPLACE versus TITRATION Protocol outcome 1: Relapse of hyperthyroidism - Actual outcome for Graves' disease: Relapse (FT4> 25pmol/l) combined with TSH <0.05 mlU/l at 24 months after ATD withdrawal; Group 1: 49/98, Group 2: 41/91 Risk of bias: All domain - High, Selection - High, Blinding - High, 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: 12, Reason: dropped out before 12 month treatment (n=11), due to pregnancy, change in residence, non-compliance, treatment with surgery or radioiodine, development of blocking TRAb, side effects of ATD, aggressive ophthalmopathy; dropped out after treatment (n=1) ; Group 2 Number missing: 17, Reason: dropped out before 12 month treatment (n=16), due to pregnancy, change in residence, non-compliance, treatment with surgery or radioiodine, development of blocking TRAb, side effects of ATD, aggressive ophthalmopathy, dropped out after treatment (n=1)	
Protocol outcomes not reported by the study	Quality of life; Mortality; Thyroid ophthalmopathy; Euthyroidism; Hypothyroidism; Ischaemic heart disease; Heart failure; Arrhythmia; Osteoporosis; Impaired cognitive function; Growth; Pain; Symptom scores; Experience of care; Healthcare contacts; Agranulocytosis; Liver failure; Minor drug related adverse events; Teratogenesis

Study	Peixoto 2006[68]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=55)
Countries and setting	Conducted in Brazil
Line of therapy	1st line
Duration of study	Intervention + follow up: 12-month intervention + 12 - 38-month follow up
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Radioimmunoassay, every two months
Stratum	Graves' disease
Subgroup analysis within study	Not applicable
Inclusion criteria	unequivocal Graves' disease (clinical signs of hyperthyroidism combined with low serum TSH plus elevated serum thyroid hormone levels and positive TRAb or ophthalmopathy)
Exclusion criteria	pregnancy, ongoing immunosuppressive therapy, noncompliance because of psychiatric disease, patient's preference for surgery or radioiodine treatment, or unwillingness to participate in the study,
Recruitment/selection of patients	Not specified
Age, gender and ethnicity	Age - Mean (SD): 37.7 (10.5). Gender (M:F): 13/42. Ethnicity: Not specified
Further population details	1. Age: 2. Gender:
Indirectness of population	No indirectness
Interventions	(n=30) Intervention 1: Carbimazole/methimazole. 40 to 60 mg daily. Duration 12 months. Concurrent medication/care: Not specified. Indirectness: No indirectness (n=25) Intervention 2: Propylthiouracil. 200 to 300 mg every 12 hours. Duration 12 months. Concurrent medication/care: Not specified. Indirectness: No indirectness
Funding	No funding
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: CARBIMAZOLE/METHIMAZOLE versus PROPYLTHIOURACIL Protocol outcome 1: Euthyroidism - Actual outcome for Graves' disease: Remission at 12 months: Group 1: 10/25, Group 2: 15/21	

Risk of bias: All domain - High, Selection - High, Blinding - High, 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: 5, Reason: Non-compliance, pregnancy; Group 2 Number missing: 4, Reason: Non-compliance, pregnancy, severe side effects Protocol outcome 2: Minor drug related adverse events - Actual outcome for Graves' disease: Minor side effects at Not specified; Group 1: 2/25, Group 2: 0/21; Comments: Dose dependent: favoring low dose Risk of bias: All domain - Very high, Selection - High, Blinding - Very high, 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: 5, Reason: Non-compliance, pregnancy; Group 2 Number missing: 4, Reason: Non-compliance, pregnancy, severe side effects	
Protocol outcomes not reported by the study	Quality of life; Mortality; Thyroid ophthalmopathy; Hypothyroidism; Relapse of hyperthyroidism; Ischaemic heart disease; Heart failure; Arrhythmia; Osteoporosis; Impaired cognitive function; Growth; Pain; Symptom scores; Experience of care; Healthcare contacts; Agranulocytosis; Liver failure; Teratogenesis

Study	Rittmaster 1998[77]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=199)
Countries and setting	Conducted in Canada; Setting:
Line of therapy	1st line
Duration of study	Follow up (post intervention): mean 27 months (range: 6-47)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: symptoms and biochemical evidence of hyperthyroidism
Stratum	Graves' disease
Subgroup analysis within study	Not applicable
Inclusion criteria	active, previously untreated Graves' disease based on symptoms of hyperthyroidism, a thyroid examination consistent with Graves' disease, biochemical evidence of hyperthyroidism and an increased thyroidal uptake of radioiodine or a rapid and diffuse uptake of technetium
Exclusion criteria	not specified
Recruitment/selection of patients	not specified
Age, gender and ethnicity	Age - Mean (SD): 38(14). Gender (M:F): 23/126. Ethnicity: Caucasian (n=144), Native American (n=3), Asian (n=1) and African American (n=1)
Further population details	1. Age: 2. Gender:
Indirectness of population	No indirectness
Interventions	(n=98) Intervention 1: Block and replace. 15 mg MMI twice daily & T4 sufficient dose to maintain TSH in the mid- to high-normal range (2.0-5.4 mIU/L) or TSH less than or equal to 0.6 mIU/L. Duration 18-months. Concurrent medication/care: 10 mg MMI three times daily for mean (SD): 7.9 (6.2) weeks, until serum total T3 concentration entered normal range (0.9-2.8 nmol/L). Indirectness: No indirectness (n=51) Intervention 2: Titration. MMI adjusted to maintain normal TSH (0.3-5.4 mIU/L). Duration 18 months. Concurrent medication/care: 10 mg MMI three times daily for mean (SD): 7.9 (6.2) weeks, until serum total T3 concentration entered normal range (0.9-2.8 nmol/L). Indirectness: No indirectness
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: BLOCK AND REPLACE versus TITRATION Protocol outcome 1: Relapse of hyperthyroidism - Actual outcome for Graves' disease: Relapse at mean 27 months after treatment withdrawal (range: 6-47 months); Group 1: 21/98, Group 2: 18/51 Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - High, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - High; Indirectness of outcome: No indirectness ; Group 1 Number missing: ; Group 2 Number missing:	
Protocol outcomes not reported by the study	Quality of life; Mortality; Thyroid ophthalmopathy; Euthyroidism; Hypothyroidism; Ischaemic heart disease; Heart failure; Arrhythmia; Osteoporosis; Impaired cognitive function; Growth; Pain; Symptom scores; Experience of care; Healthcare contacts; Agranulocytosis; Liver failure; Minor drug related adverse events; Teratogenesis

Study	**Romaldini 1983[79]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=113)
Countries and setting	Conducted in Brazil; Setting: Hospital
Line of therapy	Unclear
Duration of study	Intervention + follow up: 10 – 30-month intervention, 17 - 81 month follow-up
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: based on clinical grounds
Stratum	Graves' disease
Subgroup analysis within study	Not applicable
Inclusion criteria	patients with unequivocal Graves' hyperthyroidism, diagnosis based on clinical grounds, confirmed by the determination of serum thyroid hormone levels, thyroid autoantibodies, serum TSH levels, radioactive iodine uptake (RAIU), and scintigraphy
Exclusion criteria	not specified
Recruitment/selection of patients	not specified
Age, gender and ethnicity	Age - Mean (SD): High dose: 40(11); titrated dose: 40(13). Gender (M:F): 18/95. Ethnicity: not specified
Further population details	1. Age: 2. Gender:
Indirectness of population	No indirectness
Interventions	(n=65) Intervention 1: Block and replace. MMI 40-100 mg/d (mean (SD) 60.7 (14.5) mg, n=34) or PTU 500-1200 mg/d (mean (SD) 694 (173) mg, n=31); large start dose, increased to obtain total blockage when necessary; 50-75 µg T3 added 2-3 weeks after. Duration 10-30 months. Concurrent medication/care: not specified. Indirectness: No indirectness Comments: Drugs given at 8-hour intervals (n=48) Intervention 2: Titration. MMI 40 mg or PTU 500 mg, gradually reduced to MMI 5-25 mg/d (mean (SD):13.6(7), n=25) or PTU 100-300 mg/d (mean (SD): 180(58), n=23) to maintain euthyroid state. Duration 12-20 months. Concurrent medication/care: not specified. Indirectness: No indirectness Comments: Drugs given at 8-hour intervals
Funding	Academic or government funding (CNPq)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: BLOCK AND REPLACE versus TITRATION Protocol outcome 1: Relapse of hyperthyroidism - Actual outcome for Graves' disease: Relapse at 17-81 months after treatment discontinuation; Group 1: 16/65, Group 2: 28/48 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 - High; 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 Thyroid ophthalmopathy; Euthyroidism; Hypothyroidism; Ischaemic heart disease; Heart failure; Arrhythmia; Osteoporosis ; Impaired cognitive function ; Growth; Pain; Symptom scores; Experience of care; Healthcare contacts; Agranulocytosis; Liver failure; Minor drug related adverse events; Teratogenesis

Study	Weetman 1994[91]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=104)
Countries and setting	Conducted in United Kingdom; Setting: endocrine clinic in Cambridge
Line of therapy	1st line
Duration of study	Intervention + follow up: 6 or 12 months + 12-month follow-up
Method of assessment of guideline condition	Method of assessment /diagnosis not stated
Stratum	Graves' disease
Subgroup analysis within study	Not applicable
Inclusion criteria	patients <55 with Graves' disease, diagnosed by the presence of hyperthyroidism with a diffuse goitre, and supported by the presence of thyroglobulin/microsomal antibodies, eye signs or a family history, suppressed TSH and elevated free T4 (FT4) levels at diagnosis
Exclusion criteria	not specified
Recruitment/selection of patients	consecutive patients
Age, gender and ethnicity	Age - Other: <55 years. Gender (M:F): 12/92. Ethnicity: Caucasian (93.3%)
Further population details	1. Age: 2. Gender:
Indirectness of population	No indirectness
Interventions	(n=51) Intervention 1: 12 – 18-month treatment. carbimazole 20mg three times/day, reduced to 40 mg once/d after 4 weeks; thyroxin started at 4 weeks at 1.5 mcg/kg daily, rounded up to the nearest 25 mcg if the patient was euthyroid or deferred for 1-2 weeks if patient was still hyperthyroid. Duration 12 months. Concurrent medication/care: not specified. Indirectness: No indirectness (n=49) Intervention 2: 6 - <12-month treatment. carbimazole 20mg three times/day, reduced to 40 mg once/d after 4 weeks; thyroxin started at 4 weeks at 1.5 mcg/kg daily, rounded up to the nearest 25 mcg if the patient was euthyroid or deferred for 1-2 weeks if patient was still hyperthyroid. Duration 6 months. Concurrent medication/care: not specified. Indirectness: No indirectness
Funding	Other author(s) funded by industry (the Wellcome Trust)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: 12-18 MONTH TREATMENT versus 6-<12 MONTH TREATMENT Protocol outcome 1: Euthyroidism - Actual outcome for Graves' disease: Remission at 12 months after end of treatment; Group 1: 33/51, Group 2: 29/49 Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Baseline details are not given; Group 1 Number missing: , Reason: four failed to complete trial, one became pregnant, three moved away; Group 2 Number missing: , Reason: four failed to complete trial, one became pregnant, three moved away	
Protocol outcomes not reported by the study	Quality of life; Mortality; Thyroid ophthalmopathy; Hypothyroidism; Relapse of hyperthyroidism; Ischaemic heart disease; Heart failure; Arrhythmia; Osteoporosis; Impaired cognitive function; Growth; Pain ; Symptom scores; Experience of care; Healthcare contacts; Agranulocytosis; Liver failure; Minor drug related adverse events; Teratogenesis

Appendix E: Forest plots

E.1 Grave's disease- methimazole/carbimazole versus propylthiouracil

Figure 2: Euthyroidism (3-12 months)

Study or Subgroup	MMI Events	MMI Total	PTU Events	PTU Total	Weight	Risk Ratio M-H, Random, 95% CI
He 2004	12	15	5	15	22.1%	2.40 [1.12, 5.13]
Homsanit 2001	27	35	7	36	23.2%	3.97 [1.99, 7.90]
Nakamura 2007	176	194	54	69	29.5%	1.16 [1.02, 1.32]
Peixoto 2006	10	25	15	21	25.2%	0.56 [0.32, 0.97]
Total (95% CI)		**269**		**141**	**100.0%**	**1.51 [0.75, 3.03]**
Total events	225		81			

Heterogeneity: Tau² = 0.43; Chi² = 24.65, df = 3 (P < 0.0001); I² = 88%
Test for overall effect: Z = 1.15 (P = 0.25)

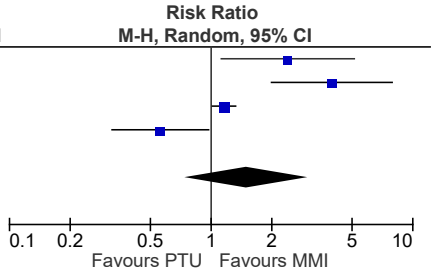


Figure 3: Hypothyroidism (12 weeks)

Study or Subgroup	Favours MMI Events	Favours MMI Total	PTU Events	PTU Total	Weight	Peto Odds Ratio Peto, Fixed, 95% CI
He 2004	4	15	0	15	27.6%	9.31 [1.17, 73.75]
Homsanit 2001	11	35	0	36	72.4%	10.66 [2.97, 38.22]
Total (95% CI)		**50**		**51**	**100.0%**	**10.27 [3.46, 30.44]**
Total events	15		0			

Heterogeneity: Chi² = 0.01, df = 1 (P = 0.91); I² = 0%
Test for overall effect: Z = 4.20 (P < 0.0001)

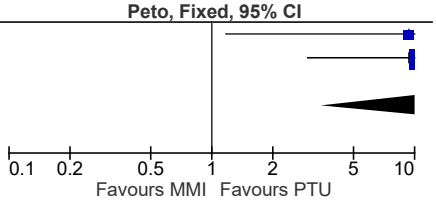
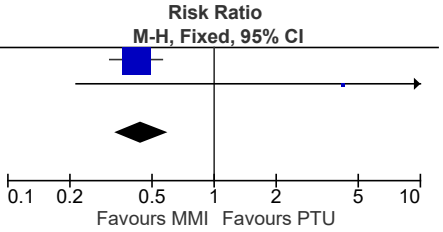


Figure 4: Minor drug related adverse events (3 – 12-months)

Study or Subgroup	MMI Events	MMI Total	PTU Events	PTU Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Nakamura 2007	58	267	54	104	99.3%	0.42 [0.31, 0.56]
Peixoto 2006	2	25	0	21	0.7%	4.23 [0.21, 83.53]
Total (95% CI)		**292**		**125**	**100.0%**	**0.44 [0.33, 0.59]**
Total events	60		54			

Heterogeneity: Chi² = 2.36, df = 1 (P = 0.12); I² = 58%
Test for overall effect: Z = 5.48 (P < 0.00001)

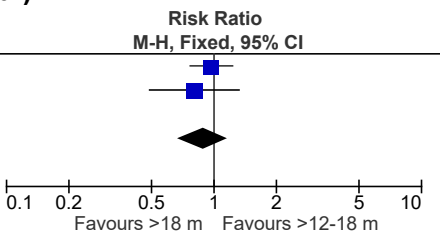


E.2 Grave's disease- 12 – 18-month vs >18-month treatment

Figure 6: Relapse (2 – 5 years post treatment withdrawal)

Study or Subgroup	>18m Events	>18m Total	12-18m Events	12-18m Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Garcia-Mayor 1992	20	24	24	28	47.9%	0.97 [0.77, 1.23]
Maugendre 1999	18	62	26	72	52.1%	0.80 [0.49, 1.32]
Total (95% CI)		**86**		**100**	**100.0%**	**0.88 [0.67, 1.16]**
Total events	38		50			

Heterogeneity: Chi² = 0.77, df = 1 (P = 0.38); I² = 0%
Test for overall effect: Z = 0.88 (P = 0.38)



E.3 Grave's disease- 6 - <12-month vs 12 – 18-month treatment

Figure 7: Relapse (24 months post treatment withdrawal)

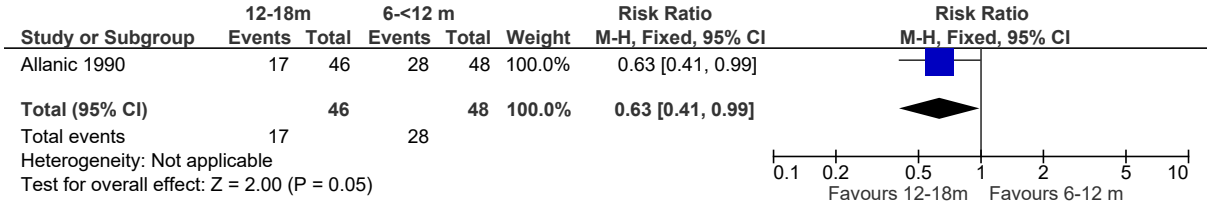
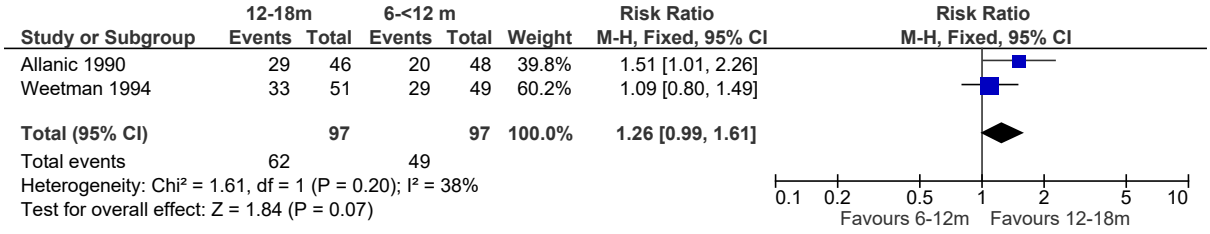


Figure 8: Euthyroidism (12 – 24 months post treatment withdrawal)



E.4 Grave's disease- block-replace versus titration

Figure 9: Relapse (up to 47 months after treatment withdrawal)

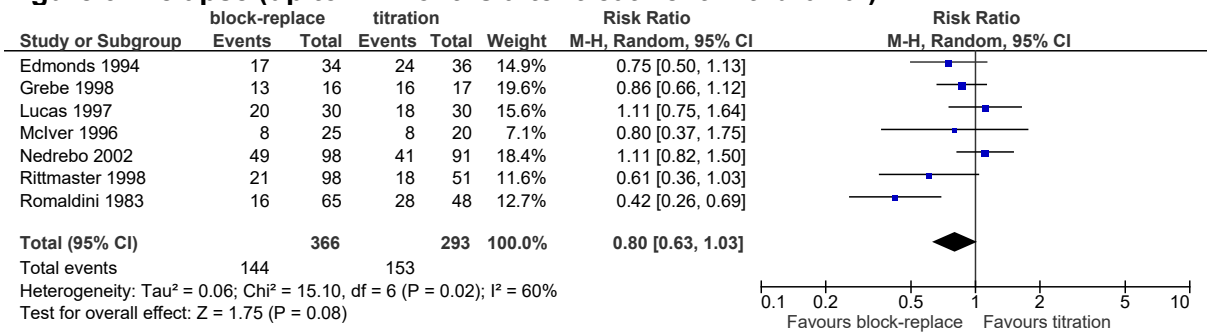


Figure 10: Minor drug related adverse events (during 6 – 12-month treatment)

Study or Subgroup	Favours block-replace Events	Total	titration Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Edmonds 1994	7	49	6	46	63.3%	1.10 [0.40, 3.02]
Grebe 1998	5	17	0	20	36.7%	12.83 [0.76, 216.55]
Total (95% CI)		**66**		**66**	**100.0%**	**2.70 [0.23, 31.79]**
Total events	12		6			

Heterogeneity: Tau² = 2.23; Chi² = 2.90, df = 1 (P = 0.09); I² = 66%
Test for overall effect: Z = 0.79 (P = 0.43)

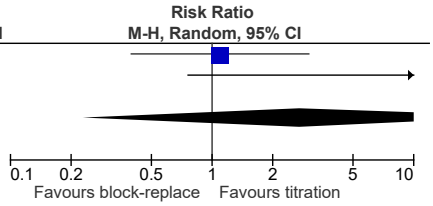
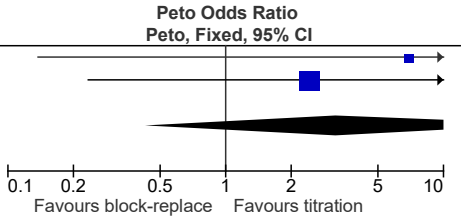


Figure 11: Agranulocytosis (during 6 – 12 month treatment)

Study or Subgroup	block-replace Events	Total	titration Events	Total	Weight	Peto Odds Ratio Peto, Fixed, 95% CI
Edmonds 1994	1	49	0	46	26.2%	6.95 [0.14, 350.96]
Grebe 1998	2	17	1	20	73.8%	2.42 [0.23, 25.02]
Total (95% CI)		**66**		**66**	**100.0%**	**3.19 [0.43, 23.74]**
Total events	3		1			

Heterogeneity: Chi² = 0.21, df = 1 (P = 0.65); I² = 0%
Test for overall effect: Z = 1.13 (P = 0.26)



Appendix F: GRADE tables

Table 9: Clinical evidence profile: MMI versus PTU

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Methimazole/carbimazole	Propylthiouracil	Relative (95% CI)	Absolute		
Euthyroidism (follow-up 3-12 months; assessed with: cases)												
4	randomised trials	serious[1]	serious[2]	no serious indirectness	serious[3]	none	225/269 (83.6%)	52.4%	RR 1.51 (0.75 to 3.03)	267 more per 1000 (from 131 fewer to 1000 more)	⊕OOO VERY LOW	IMPORTANT
Hypothyroidism (follow-up 12 weeks; assessed with: cases)												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[3]	none	15/50 (30%)	0%	Peto OR 10.27 (3.46 to 30.44)	300 more per 1000 (from 0 more to 0 more)[4]	⊕OOO VERY LOW	IMPORTANT
Minor drug related adverse events (follow-up 3-12 months; assessed with: cases)												
2	randomised trials	very serious[1]	serious[2]	no serious indirectness	no serious imprecision	none	60/292 (20.5%)	54/125 (43.2%)	RR 0.44 (0.33 to 0.59)	146 fewer per 1000 (from 107 fewer to 174 fewer)	⊕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 or 2 increments because the point estimate and or the confidence intervals varied widely across studies, unexplained by subgroup analysis
[3] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs
[4] Zero events in control group

Table 8: Clinical evidence profile: 12 – 18-month vs > 18-month treatment

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	>18m	12-18m	Relative (95% CI)	Absolute		
Relapse (follow-up 2-5 years; assessed with: cases (post treatment withdrawal))												
2	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	Serious[2]	none	38/86 (44.2%)	60.9%	RR 0.88 (0.67 to 1.16)	73 fewer per 1000 (from 201 fewer to 97 more)	⊕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

Table 9: Clinical evidence profile: 6-<12-month vs 12 - 18-month treatment

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	12-18m	6-<12m	Relative (95% CI)	Absolute		
Relapse (follow-up 24 months; assessed with: cases (post treatment withdrawal))												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	17/46 (37%)	58.3%	RR 0.63 (0.41 to 0.99)	216 fewer per 1000 (from 6 fewer to 344 fewer)	⊕⊕⊕O MODERATE	IMPORTANT
Euthyroidism (follow-up 12-24 months; assessed with: cases (post treatment withdrawal))												
2	randomised trials	Serious[2]	no serious inconsistency	no serious indirectness	serious[1]	none	62/97 (63.9%)	50.4%	RR 1.26 (0.99 to 1.61)	131 more per 1000 (from 5 fewer to 307 more)	⊕⊕OO LOW	IMPORTANT

[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

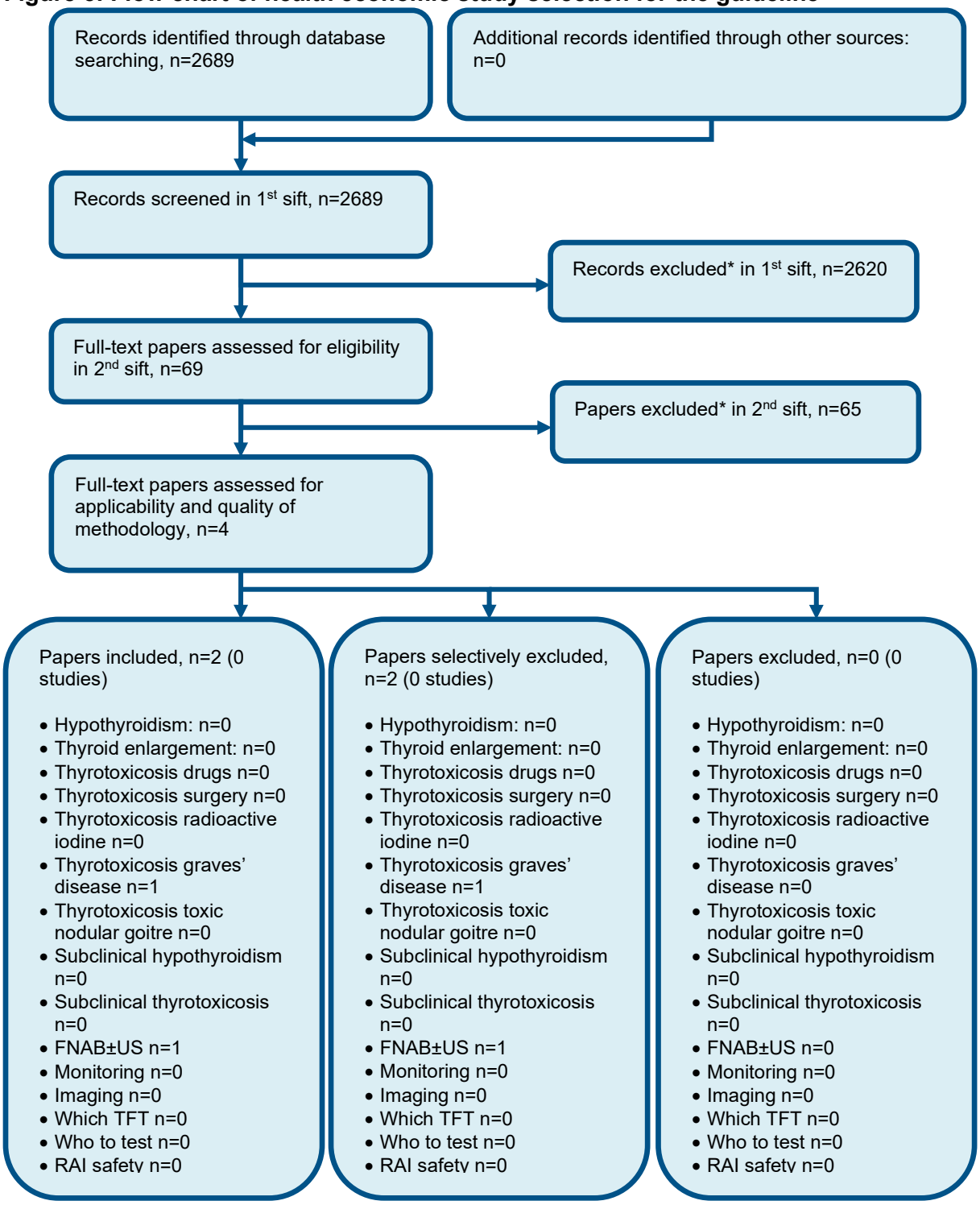
Table 10: Clinical evidence profile: Block-replace versus titration

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Block-replace	Titration	Relative (95% CI)	Absolute		
Relapse (follow-up 6-47 months; assessed with: cases (post treatment withdrawal))												
7	randomised trials	serious[1]	serious[2]	no serious indirectness	serious[3]	none	144/366 (39.3%)	58.3%	RR 0.8 (0.63 to 1.03)	117 fewer per 1000 (from 216 fewer to 17 more)	⊕OOO VERY LOW	IMPORTANT
minor drug related adverse events (follow-up 6-12 months; assessed with: cases (during treatment))												
2	randomised trials	serious[1]	Serious[2]	no serious indirectness	serious[3]	none	12/66 (18.2%)	6.5%	RR 2.70 (0.23 to 31.79)	59 more per 1000 (from 50 fewer to 935 more)	⊕OOO VERY LOW	IMPORTANT
Agranulocytosis (follow-up 6-12 months; assessed with: cases (during treatment))												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	Serious[3]	none	3/66 (4.5%)	2.5%	Peto OR 3.19 (0.43 to 23.74)	51 more per 1000 (from 14 fewer to 353 more)	⊕⊕OO 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 or 2 increments because the point estimate and or the confidence interval varied widely across studies, unexplained by subgroup analysis.
[3] 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 5: Flow chart of health economic study selection for the guideline



* Non-relevant population, intervention, comparison, design or setting; non-English language

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
Abraham-nordling 2007[1]	No usable outcomes
Andrade 1999[4]	Less than minimum duration
Andrade 2001[5]	Incorrect interventions
Andrade 2004[6]	Incorrect interventions
Azizi 2012[8]	Wrong study design
Azizi 2018 [7]	NRS where RCTs are available
Barczynski 2012[9]	Incorrect interventions
Barczynski 2010[10]	Abstract only
Barczynski 2018[11]	Incorrect interventions
Benker 1995[13]	Incorrect interventions
Benker 1998[12]	Incorrect interventions
Bonnema 2003[15]	Incorrect interventions
Bonnema 2004[16]	Incorrect interventions
Bonnema 2011[17]	Inappropriate comparison
Braga 2002[18]	Less than minimum duration
Burch 2001[19]	No usable outcomes
Buscemi 2007[20]	Not guideline condition
Canto 2016[21]	Incorrect interventions
Chen 2011[22]	Inappropriate comparison
Chen 2014[23]	No additional outcomes to those reported elsewhere
Chi 2005[24]	Inappropriate comparison
Connell 1987[25]	No usable outcomes
De Luca 2018[26]	SR, checked for references
Esfahani 2005[28]	Inappropriate comparison
Glinoer 2001[30]	Incorrect interventions
Goni iriarte 1995[31]	Not in English
Hamide 2014[33]	NRS where RCTs are available
Hashizume 1991[34]	NRS without adequate adjustment
Hoermann 2002[36]	Incorrect interventions
Howarth 2001[38]	Incorrect interventions
Jaiswal 2014[39]	Incorrect interventions
Järhult 2005[40]	Incorrect interventions
Jorde 1995[42]	Incorrect interventions
Kallner 1996[43]	Incorrect interventions
Kung 1995[44]	Incorrect interventions
Leclere 1994[45]	Not in English
Leslie 2003[46]	Incorrect interventions
Leung 2017 [47]	SR, checked for references
Li 2016[48]	SR, checked for references
Liu 2015[50]	Incorrect interventions

Study	Exclusion reason
Liu 2017[49]	Incorrect interventions
Ljunggren 1998[51]	No usable outcomes
Ma 2008[53]	SR, checked for references
Ma 2016[54]	SR checked for references
Marcocci 1989[55]	Incorrect interventions
Mashio 1997[56]	Inappropriate comparison
Mastorakos 2003[57]	Incorrect interventions
Menconi 2007[60]	No usable outcomes
Miranda-padua 2014[61]	Incorrect interventions
Müller 2001[62]	Inappropriate comparison
Noh 2015[66]	Incorrect interventions
Orsini 2012[67]	Inappropriate comparison
Peters 1995[69]	Incorrect interventions
Peters 1996[70]	No usable outcomes
Peters 1997[71]	Incorrect interventions
Pfeilschifter 1997[72]	Inappropriate comparison
Pirnat 2011[73]	Incorrect interventions
Pusuwan 2011[74]	Inappropriate comparison
Raber 2000[75]	Incorrect interventions
Reinwein 1993[76]	Inappropriate comparison
Rokni 2014[78]	SR checked for references
Santos 2004[80]	NRS without adequate adjustment
Santos 2012[81]	Inappropriate comparison
Sapienza 2015[82]	Inappropriate comparison
Schneider 2005[83]	Inappropriate comparison
Singhal 2014[84]	Withdrawn Cochrane review
Taïeb 2016[85]	Incorrect interventions
Thientunyakit 2010[86]	Inappropriate comparison
Tian 2001[87]	Not in English
Unalp 2009[88]	No usable outcomes
Walter 2006[89]	NRS without adequate adjustment
Wang 2016[90]	SR, checked for references
Witte 2000[92]	Incorrect interventions
Yousefi 2011[93]	Not in English
Yuan 2017[94]	SR, checked for references

J.2 Excluded health economic studies

None

Appendix K: Research recommendations

K.1 Research question: What is the clinical and cost effectiveness of a block and replace regimen compared with a titration regimen of antithyroid drugs for Graves' disease?

Why this is important:

Antithyroid drugs (ATDs) are a commonly used treatment modality for Graves' hyperthyroidism/disease. There are two regimes of ATDs: (a) 'block and replace' regime (a fixed high dose of ATD is combined with levothyroxine) and (b) 'titration' regime (titrated dose of ATD based on thyroid function tests). It remains uncertain which of these two regimes is most effective for treating Graves' hyperthyroidism in terms of remission rate, adverse effects and stability of thyroid function. Limitations in the current evidence have led to conflicting recommendations in the international guidelines for the management of Graves' hyperthyroidism and to variation in clinical practice. A national survey showed that one third of UK endocrinologists use a block and replace regime while the others prefer a titration regime.

The evidence currently identified was of low quality and was thus insufficient to allow us to draw conclusions between these two options. A large high quality trial comparing the clinical and cost effectiveness of these two regimes for people with Graves' disease will help to reduce the variation in clinical practice and improve patient care.

Criteria for selecting high-priority research recommendations:

PICO question	Population: People with Graves' hyperthyroidism/disease who are being treated with an antithyroid drug (ATD) Intervention(s): Block and replace regime of ATD Comparison: Titration regime of ATD Outcome(s): quality of life, symptom control, biochemical euthyroidism, side effects of ATD, new development and worsening of thyroid eye disease, hyperthyroidism relapse rate, cost
Importance to patients or the population	This research will help to establish which of the two regimes of ATDs is most clinically and cost-effective, leading to reduction in variation in clinical practice and improvement in patient care.
Relevance to NICE guidance	This research will enable future guidelines to identify and recommend the most clinically and cost effective regime for treating people with Graves' hyperthyroidism.
Relevance to the NHS	Clear evidence supporting the superiority of one ATD regime over the other in terms of clinical and cost-effectiveness will offer clinicians clear guidance on the preferred ATD regime for the management of people with Graves' disease.
National priorities	Hyperthyroidism, most frequently caused by Graves' disease, comes under the long-term condition directorate in the UK.
Current evidence	Several randomised controlled trials (RCTs) and retrospective

base	observational studies have compared clinical outcomes with the block and replace regime versus the titration regime for treatment of Graves' disease. However, as most of these studies are small and associated with methodological limitations, it is difficult to derive a firm conclusion regarding the effectiveness of any of the two options over the other. A Cochrane review on this topic has highlighted the need for further research.
Equality	This recommendation will help to reduce the current variation in clinical practice in the UK.
Study design	RCTwith corresponding economic analysis.
Feasibility	Considering the wide administration of ATDs under both regimes across the UK, a multi-centre UK trial is feasible.
Other comments	
Importance	Medium: the research is relevant to the recommendations in the guideline, but the research recommendation is not key to future updates.

K.2 Research question: What is the clinical and cost effectiveness of different durations of antithyroid drug regimens for people with T3 thyrotoxicosis due to Graves' disease?

Why this is important:

T3 thyrotoxicosis is the mildest form of overt hyperthyroidism but currently patients are treated in the same way as any person with Graves' disease, using antithyroid drugs for 12-18 months. This is largely because the randomised trials using antithyroid drugs to treat Graves' disease were performed 20-30 years ago, before T3 thyrotoxicosis could be identified reliably by biochemical testing. Thus, it may be the case that this patient group, which accounts for 20-25% of patients currently presenting with Graves' disease, is being unnecessarily exposed to prolonged antithyroid drug treatment and to its associated risk of serious side-effects and excess resource use, There is no strong evidence to guide treatment for the subgroup of people with T3 thyrotoxicosis and a randomised study would clarify whether a shorter and lower dose antithyroid drugs regimen would lead to more clinically and cost-effective treatment for those people

Criteria for selecting high-priority research recommendations:

PICO question	Population: Adult patients with new onset T3 thyrotoxicosis due to Graves' disease (TRAb positive, TSH <0.05, FT3 6.5-10pmol/l, normal FT4) Intervention(s): Carbimazole 5mg daily until serum TSH in reference range on 2 consecutive readings 6 weeks apart (or ?>1.0mIU/l once) Comparison: Conventional carbimazole dose (20mg, then tapering) for 12 months Outcome(s): Proportion of patients remaining euthyroid (TSH in reference

	range) one year after withdrawal of antithyroid drugs
Importance to patients or the population	A clinical trial determining the clinical and cost-effectiveness of different (shorter) antithyroid drug regimen durations for people with T3 thyrotoxicosis could help improve treatment outcomes and minimise side effects for those people with a positive resource impact.
Relevance to NICE guidance	This will address the lack of evidence available to guide the management of people with T3 thyrotoxicosis. There is currently no distinction in management of patients with Graves' hyperthyroidism and those with T3 thyrotoxicosis in NICE guidance because there is no evidence to guide a different management approach. New knowledge could lead to a safer, cheaper stratified approach.
Relevance to the NHS	Evidence of the clinical and cost-effectiveness of different antithyroid drug regimens of shorter duration and lower dose would ensure improved patient outcomes and less resource use, in terms of drugs, clinic time, patient safety and follow up
National priorities	Efficient health resource use
Current evidence base	The problem with the current evidence base is that this patient group had too mild a disease to be included in the randomised studies which form the basis of that current evidence base. This is historical, as free T3 (FT3) assays only started to become reliable about 20 years ago, after these studies were completed. Prior to that, insensitive total T3 (TT3) assays were used which failed to identify the majority of patients with T3-thyrotoxicosis.
Equality	The majority of people with Graves' disease are women (6:1), so overtreatment currently affects mostly women.
Study design	Primary research: randomised study, powered to find 'non-inferiority' of short-term low-dose treatment.
Feasibility	Feasible within a multicentre NHS environment
Other comments	No previous attempts to answer this question have been made.
Importance	Medium: the research is relevant to the recommendations in the guideline, but the research recommendations are not key to future updates. It is widely acknowledged by experts that there is likely to be unnecessary overtreatment here which is not good medicine or health policy.