National Institute for Health and Care Excellence

FINAL

Thyroid disease: assessment and management

[I] Management of thyrotoxicosis:
- drugs vs surgery vs radioactive iodine
- safety of treatment with radioactive iodine

NICE guideline NG145

Intervention evidence review underpinning recommendations 1.6.7 to 1.6.26 in the guideline. See also evidence reviews J, 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 All NICE guidance is subject to regular review and may be updated or withdrawn.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

1 Management of thyrotoxicosis: drugs vs surgery vs radioactive iodine

1.1 Review question: What is the clinical and cost effectiveness of using radioactive iodine vs antithyroid drugs vs surgery to treat thyrotoxicosis secondary to Graves' disease?

Review question: What is the clinical and cost effectiveness of using radioactive iodine vs surgery to treat thyrotoxicosis secondary to toxic nodular goitre?

1.2 Introduction

The three principal treatment modalities when managing the patient with thyrotoxicosis are medical therapy with antithyroid drugs (ATD), radioactive iodine or surgery. There is uncertainty in terms of how these modalities are best used in relation to the type and dose of antithyroid drugs, the dose of radioactive iodine and the nature of the surgical procedure (partial or total thyroidectomy). The aetiology of thyrotoxicosis (such as Graves' disease, toxic nodular goitre, toxic nodule or thyroiditis), the age of the patient, other patient factors (such as pregnancy or planned pregnancy and small children at home) and the presence of complicating factors such as thyroid eye disease are additional considerations. Although many patients with Graves' thyrotoxicosis are managed with ATD (carbimazole or propylthiouracil) initially, a majority will relapse and become thyrotoxic again when the drugs are stopped. Patients will then be faced with the prospect of long term ATD therapy or choosing radioactive iodine or surgery, both of which can potentially result in hypothyroidism and a requirement for life-long thyroid hormone replacement.

Radioactive iodine has been used to treat thyrotoxicosis for many years. The attractions of this therapy include the fact that it is relatively cheap. Administration is straight-forward although guidelines that limit exposure to ionising radiation need to be followed when using radioactive agents and there is variation between centres in terms of when this modality is considered to be an appropriate therapeutic option. Some units are more proactive than others and consider this treatment more readily in the context of the younger patient and the individual with complicating factors such as thyroid eye disease. Establishing the circumstances and threshold for using this treatment is an important area because the therapeutic options for patients who fail to respond to anti-thyroid drugs are limited.

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
Interventions	Antithyroid drugs Radioactive iodine Surgery
Comparison	Any of the above compared with any other
Outcomes	**Critical**

	• Mortality • Quality of life **Important (general)** • Thyroid ophthalmopathy • Euthyroidism • Hypothyroidism • Relapse of hyperthyroidism • Cardiovascular morbidity • Arrhythmia • Osteoporosis • Cognitive impairment • Pain • Symptom scores • Patient/family/carer experience • Healthcare contacts **Important (surgical)** • Recurrent laryngeal nerve damage • Hypocalcaemia • Hypoparathyroidism • Bleeding • Infection **Important (pharmacological)** • Agranulocytosis • Liver failure • Minor drug related adverse effects • Teratogenesis **Important (radioactive iodine)** • Infertility • Malignancy • Thyrotoxic storm • Growth abnormalities • Hypocalcaemia • Hypoparathyroidism • Teratogenesis
Study design	RCTs only, non-randomised studies only if key confounders (age, co-existing conditions, baseline thyroid hormones) taken into account Minimum duration 3 months

1.4 Clinical evidence

1.4.1 Included studies

Six randomised controlled studies (in nine publications) were included in the review;[1, 3, 11, 16, 28, 70, 124, 127, 128] 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,[81] the studies included in this review were checked against the protocol and included as appropriate.

Five studies compared antithyroid drugs vs radioactive iodine. One study compared antithyroid drugs vs radioactive iodine vs surgery.

Five studies were in the treatment naïve population (or previous treatment unspecified). One study was in people who had previously used antithyroid drugs and relapsed.

No studies were found in children or older adults.

All six studies were either exclusively in people with Graves' disease or in a mixed population in which the majority had Graves' disease. No studies were found exclusively in people with toxic nodular goitre.

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
Azizi 2005[11]	Antithyroid drugs, n = 52 MMI, 10mg twice daily for first month, once daily for second month, 2.5-10mg daily thereafter, no discontinuation specified Radioactive iodine , n = 52 Calculated activity, no information on number of treatments	Adults (mean age 48, SD 6) Second line (relapsed 1 year after 18 months of antithyroid drug use) Graves' disease Percent with ophthalmopathy at baseline not specified Iran	Euthyroidism (at end of follow-up) Hypothyroidism (at end of follow-up) Hyperthyroidism (at end of follow-up) Agranulocytosis 10 year follow-up	In all patients dosage of MMI/levothyroxine adjusted to maintain normal thyroid function No discontinuation period specified for ATDs
Bartalena 1998[16]	Antithyroid drugs, n = 148 Lowest dose that maintained euthyroidism, no discontinuation specified Radioactive iodine , n = 150 MMI given for 3 to 4 months prior to RAI, stopped 5 days before, dose of 120-150uCi per gram of thyroid tissue, if hypo/hyperthyroid after RAI corrected with levothyroxine or MMI as relevant, second dose of RAI at end of follow-up if persistent hyperthyroidism still	Adults (mean age 42, range 15-85) 70% had received MMI prior to referral Graves' disease ~50% with ophthalmopathy (mild) at baseline Italy	Ophthalmopathy (new/worsening) Euthyroidism (at end of follow-up) Hypothyroidism (at end of follow-up) Hyperthyroidism (at end of follow-up) 1 year follow-up (no discontinuation period for ATDs)	Radioactive iodine arm given levothyroxine/MMI if required No discontinuation period for ATDs
Chen 2009[28]	Antithyroid drugs, n = 230	Adults (mean age 37, SD 14)	Mortality	No information on drug

Study	Intervention and comparison	Population	Outcomes	Comments
	Either MMI or PTU, at least 18 months of treatment, dose based on severity of symptoms and TSH Radioactive iodine , n = 230 Calculated activity, no pre-treatment with ATDs, at 3 months could have 2nd treatment (10% participants) an at 6 months could have 3rd (2.5% participants)	Treatment naïve Mixed cause (75% Graves', 25% toxic nodular goitre) ~25% with ophthalmopathy at baseline China	Ophthalmopathy (incidence) Euthyroidism (at end of follow-up) Hypothyroidism (at end of follow-up) Hyperthyroidism Agranulocytosis Severe liver damage Malignancy Thyroid storm 9 year follow-up	supplementation of radioactive iodine arm Discontinuation period for ATDs
Kansara 2017[70]	Antithyroid drugs, n = 30 CZL, 30mg initially and then tapered Radioactive iodine , n = 30 Single oral dose of 10mCi, no stated pre-treatment with ATDs	Adults (mean age 33, SD 4.2) Treatment naïve Mixed cause (85% Graves', 15% toxic nodular goitre) Percent with ophthalmopathy at baseline not specified India	Euthyroidism (at end of follow-up) Hypothyroidism (at end of follow-up) Hyperthyroidism (at end of follow-up) 1 year follow-up	No discontinuation period for ATDs
Torring 1996[1, 124, 127]	Antithyroid drugs, n = 71 MMI, 18 months of treatment, block and replace Radioactive iodine , n = 39 Oral dose, calculated activity,	Adults (younger adults mean age 29, SD 4, older adults mean 45, SD 6) Previous treatment not specified	Ophthalmopathy (new/worsening) Hyperthyroidism Recurrent laryngeal nerve damage Hypoparathyroidism	Study stratified by age group, older adults (35-55) randomised to all 3 treatments, younger adults (20-34) only randomised to antithyroid drugs or surgery. Evidence combined across age

Study	Intervention and comparison	Population	Outcomes	Comments
	(~50% of participants required more than one dose, given >10 weeks after first) Surgery, n = 64 Bilateral subtotal thyroidectomy with posterior capsule and 1g or less of each lobe left behind, thyroxine after surgery	Graves' disease ~13% with ophthalmopathy at baseline (non-severe) Sweden	Agranulocytosis Maximum 21 year follow-up	groups as per protocol, except where this would affect group age composition (i.e. not comparing older adults receiving radioactive iodine with mix of young and old adults receiving antithyroid drugs) Discontinuation of ATDs
Träisk 2009[3, 128]	Antithyroid drugs, n = 150 MMI, 18 months of treatment, block and replace Radioactive iodine , n = 163 Oral outpatient dose, calculated activity, no information on number of doses, levothyroxine substitution as required, no prophylactic steroid use	Adults (mean 51, SD 8) Treatment naïve Graves' disease ~13% with ophthalmopathy at baseline (non-severe) Sweden	Ophthalmopathy (new/worsening) Hyperthyroidism (relapse) 4 years follow-up	Radioactive iodine arm given levothyroxine if required Discontinuation of ATDs

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: Radioactive iodine vs antithyroid drugs, Graves' disease, first line treatment

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with ATD	Risk difference with RAI (95% CI)
Mortality	386 (1 study)	⨁⨁⊖⊖ LOW1,2	Not estimable	0 per 1000	not estimable[5]

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with ATD	Risk difference with RAI (95% CI)
	9 years	due to risk of bias, imprecision			
Ophthalmopathy (new/worsening cases)	948 (4 studies) 1-9 years	⊕⊕⊕⊖ MODERATE1 due to risk of bias	RR 2.17 (1.64 to 2.88)	103 per 1000	121 more per 1000 (from 66 more to 194 more)
Euthyroidism (at end of follow-up)	741 (3 studies) 1-9 years	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, inconsistency	RR 0.78 (0.37 to 1.62)	759 per 1000	167 fewer per 1000 (from 478 fewer to 471 more)
Hypothyroidism (at end of follow-up)	741 (3 studies) 1-9 years	⊕⊕⊕⊖ MODERATE1 due to risk of bias	RR 5.89 (3.12 to 11.11)	34 per 1000	166 more per 1000 (from 72 more to 344 more)
Hyperthyroidism (persistence/recurrence)	1102 (5 studies) 1-9 years	⊕⊕⊖⊖ LOW1,3 due to risk of bias, inconsistency	RR 0.25 (0.09 to 0.69)	241 per 1000	181 fewer per 1000 (from 75 fewer to 219 fewer)
Osteoporosis	70 (1 study) 14-21 years	⊕⊖⊖⊖ VERY LOW1,4 due to risk of bias, imprecision	RR 1.27 (0.43 to 3.78)	139 per 1000	38 more per 1000 (from 79 fewer to 386 more)
Agranulocytosis	423 (1 study) 9 years	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	Peto OR 0.13 (0.03 to 0.6)	33 per 1000	29 fewer per 1000 (from 13 fewer to 32 fewer)
Severe liver damage	423 (1 study) 9 years	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision	Peto OR 0.14 (0.02 to 0.79)	23 per 1000	20 fewer per 1000 (from 5 fewer to 23 fewer)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with ATD	Risk difference with RAI (95% CI)
	(1 study) 9 years	LOW1,2 due to risk of bias, imprecision	estimable	0 per 1000	not estimable[5]
Thyroid storm	386 (1 study) 9 years	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	Not estimable	0 per 1000	not estimable[5]

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 as zero events in at least one arm
3 Downgraded by 1 or 2 increments because the point estimate and or the confidence intervals varied widely across studies, unexplained by subgroup analysis
4 Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs
5 Zero events in both arms

Table 4: Clinical evidence summary: Surgery vs antithyroid drugs, Graves' disease, first line treatment

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with ATD	Risk difference with SUR (95% CI)
Ophthalmopathy (new/worsening cases)	129 (1 study) 4 years	⊕⊕⊖⊖ LOW1 due to imprecision	RR 1.14 (0.47 to 2.78)	123 per 1000	17 more per 1000 (from 65 fewer to 219 more)
Osteoporosis	111 (1 study) 14-21 years	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 1.57 (0.55 to 4.51)	91 per 1000	52 more per 1000 (from 41 fewer to 319 more)
Hyperthyroidism (persistence/recurrence)	133 (1 study) 4 years	⊕⊕⊕⊕ HIGH	RR 0.16 (0.06 to 0.44)	382 per 1000	321 fewer per 1000 (from 214 fewer to 359 fewer)

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

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with ATD	Risk difference with SUR (95% CI)
at very high risk of bias					

Table 5: Clinical evidence summary: Radioactive iodine vs surgery, Graves' disease, first line treatment

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with SUR	Risk difference with RAI (95% CI)
Ophthalmopathy (new/worsening cases)	76 (1 study) 4 years	⨁⨁⨁⨀ MODERATE1 due to imprecision	RR 2.06 (0.87 to 4.84)	162 per 1000	172 more per 1000 (from 21 fewer to 622 more)
Osteoporosis	68 (1 study) 14-21 years	⨁⨀⨀⨀ VERY LOW1,2 due to risk of bias, imprecision	RR 0.86 (0.32 to 2.29)	206 per 1000	29 fewer per 1000 (from 140 fewer to 266 more)
Hyperthyroidism (persistence/recurrence)	76 (1 study) 4 years	⨁⨁⨀⨀ LOW1 due to imprecision	RR 2.53 (0.73 to 8.82)	81 per 1000	124 more per 1000 (from 22 fewer to 633 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: Radioactive iodine vs antithyroid drugs, Graves' disease, second line treatment

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

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with ATD	Risk difference with RAI (95% CI)
(at end of follow-up)	(1 study) 10 years	LOW1 due to risk of bias	(0.28 to 0.62)	929 per 1000	539 fewer per 1000 (from 353 fewer to 669 fewer)
Hypothyroidism (at end of follow-up)	69 (1 study) 10 years	⨁⨁⊖⊖ LOW1 due to risk of bias	RR 17.07 (2.45 to 118.83)	36 per 1000	579 more per 1000 (from 52 more to 1000 more)
Hyperthyroidism (at end of follow-up)	69 (1 study) 10 years	⨁⊖⊖⊖ VERY LOW1 due to risk of bias, imprecision	Peto OR 0.09 (0 to 4.6)	36 per 1000	33 fewer per 1000 (from 36 fewer to 111 more)
Agranulocytosis	69 (1 study) 10 years	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	Not estimable		Not estimable[3]

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 as at least one arm with zero events
3 Zero events in both arms

See Appendix F: for full GRADE tables.

1.5 Economic evidence

1.5.1 Included studies

Management of thyrotoxicosis secondary to Graves' disease

One health economic study was identified with the relevant comparison and has been included in this review [35].This is summarised in the health economic evidence profile below (Table 7) and the health economic evidence tables in Appendix H:

Management of thyrotoxicosis secondary to toxic nodular goitre

No relevant health economic studies were identified.

1.5.2 Excluded studies

One health economic study that was relevant to this question was excluded due to assessment of limited applicability.

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

1.5.3 Summary of studies included in the economic evidence review

Table 7: Health economic evidence profile: Radioactive iodine vs anti-thyroid drugs vs surgery for Graves' disease

Study	Applicability	Limitations	Other comments	Incremental cost[(a)]	Incremental effects[(a)]	Cost effectiveness	Uncertainty
Donovan, 2016 [35] (UK and Australia) [(a)]	Directly applicable [(b)]	Minor limitations [(c)]	• Cost utility analysis • Life time horizon • Patients received either; 1. Radioactive iodine (RAI) 2. Anti-thyroid drugs (ATD) 3. Total thyroidectomy (TT).	Mean per patient: (2−1):£11,441 (3−1): £1,690 (3−2): saves £9,751	Mean per patient: (2−1): 0.44 QALYs (3−1): −0.8 QALYs (3−2): −1.24 QALYs	RAI dominated TT (less costly and more effective). ATD was not cost effective compared to RAI at the £20,000 threshold. (ICER for ATD vs RAI = £26,279 per QALY gained)	RAI was dominant over TT in all sensitivity analysis of all parameters assessed. ATD was a cost-effective alternative to RAI at the £30,000 threshold (ICER: £26,279 per QALY gained).

Abbreviations: ICER: incremental cost-effectiveness ratio; QALY: quality-adjusted life years
(a) The results presented are those of the UK analysis only.
(b) No downgrading for applicability.
(c) The estimates of relative treatment effects are not based on met-analysis of all the available evidence. Some costs have been based on the national tariff and maybe overestimated. The model has not been run probabilistically, to adequately assess parameter uncertainty.

1.5.4 Health economic modelling

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

1.5.5 Resource costs

Relevant unit costs are provided below to aid consideration of cost effectiveness for the management of thyrotoxicosis secondary to toxic nodular goitre.

Table 8: UK costs of thyroid surgery and radioactive iodine

Intervention	Unit cost	
Surgery (Thyroid Procedures with CC Score 0-4+)(a)	£3,689	
Radioactive iodine fixed dose(b)	£286.32	
Radioactive iodine calculated dose (c)	Procedures (pre and post therapy)	Unit costs
	Uptake measurement with probe ~15 mins Band 7	£10 (d)
	USS for volume calculation	£62
	Calculations, verification, report: ~ 3 hours Band 7	£75
	Total additional cost to fixed dose	£167 (e)

Source: NHS reference costs 2016-17, total HRG schedule [34].

(a) Weighted average of all 3 combined thyroid procedures with CC scores 0-1, 2-3, 4+(KA09C, KA09D, KA09E) including excess bed days with an average length of stay of 1.6 days

(b) Cost of oral delivery of radiotherapy for thyroid ablation, cost code RN51Z

(c) Estimation obtained from committee specialists

(d) Ideally allow for 3 uptake measurements (adding another £20), practice varies

(e) Total cost = £453.32 Economic considerations: trade-off between net clinical effects and costs

1.6 Evidence statements

1.6.1 Clinical evidence statements

1.6.1.1 Radioactive iodine vs antithyroid drugs, Graves' disease, first line treatment

No clinically important difference was identified for mortality (1 study, low quality), osteoporosis (1 study, very low quality), agranulocytosis (1 study, low quality), severe liver damage (1 study, low quality), malignancy (1 study, low quality), thyroid storm (1 study, low quality).

There was a clinically important benefit of radioactive iodine for hyperthyroidism (5 studies, low quality).

There was a clinically important harm of radioactive iodine for ophthalmopathy (4 studies, moderate quality), euthyroidism (3 studies, very low quality) and hypothyroidism (3 studies, moderate quality).

No evidence was identified for other outcomes.

1.6.1.2 Surgery vs antithyroid drugs, Graves' disease, first line treatment

No clinically important difference was identified for ophthalmopathy (1 study, low quality), osteoporosis (1 study, very low quality).

There was a clinically important benefit of surgery for hyperthyroidism (1 study, high quality).

No evidence was identified for other outcomes.

1.6.1.3 Radioactive iodine vs surgery, Graves' disease, first line treatment

No clinically important difference was identified for osteoporosis (1 study, very low quality).

There was a clinically important harm of radioactive iodine for ophthalmopathy (1 study, moderate quality) and hyperthyroidism (1 study, low quality).

No evidence was identified for other outcomes.

1.6.1.4 Radioactive iodine vs antithyroid drugs, Graves' disease, second line treatment

No clinically important difference was identified for hyperthyroidism (1 study, very low quality), agranulocytosis (1 study, very low quality).

There was a clinically important harm of radioactive iodine for euthyroidism (1 study, low quality) and hypothyroidism (1 study, low quality).

No evidence was identified for other outcomes.

1.6.2 Health economic evidence statements

Management of thyrotoxicosis secondary to Graves' disease

One cost–utility analysis found that anti-thyroid drugs were not cost effective at a threshold of £20,000 per QALY, compared to radioactive iodine for treating thyrotoxicosis secondary to Graves' disease (ICER: £26,279 per QALY gained compared to radioactive iodine). It also found that radioactive iodine was dominant (less costly and more effective) compared to total thyroidectomy. This analysis was assessed as directly applicable with minor limitations.

Management of thyrotoxicosis secondary to toxic nodular goitre

- No relevant economic evaluations were identified.

2 Radioactive iodine safety

2.1 Review question: What are the long term adverse events of radioactive iodine treatment for thyrotoxicosis?

2.2 Introduction

Radioactive iodine has been used to treat thyrotoxicosis for many years. The attractions of this therapy include the fact that it is relatively cheap. Administration is straight-forward although guidelines that limit exposure to ionising radiation need to be followed when using radioactive agents and there is variation between centres in terms of when this modality is considered to be an appropriate therapeutic option. There are concerns about the potential long-term risk of developing cancer because of exposure to radiation and the impact of radiation on fertility. The purpose of his review is to establish the level of risk radiation on these outcomes.

2.3 PICO table

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

Table 9: PICO characteristics of review question

Population	People being treated with radioactive iodine for thyrotoxicosis
Intervention	Radioactive iodine
Comparisons	Antithyroid drug treatment of thyrotoxicosis Surgical treatment of thyrotoxicosis Healthy controls
Outcomes	Cancer • Overall diagnoses • Diagnoses in organs that take up iodine (e.g. thyroid, small bowel) • Diagnoses in organs that do not take up iodine • Infertility
Study design	Only studies with follow-up >5 years and sample size >1000 (for adults) will be included Evidence will be considered according to the following hierarchy: • Comparative studies with hyperthyroid controls and adequate adjustment for key confounders (age, smoking) • Comparative studies with hyperthyroid controls and without adequate adjustment for key confounders • Comparative studies with healthy controls and adequate adjustment for key confounders (age, smoking) • Comparative studies with healthy controls without adequate adjustment for key confounders

2.4 Clinical evidence

2.4.1 Included studies

Eight studies were included in the review;[38, 40, 42, 44, 59, 62, 90, 115] these are summarised below. Evidence from these studies is summarised in the clinical evidence summary below (Table

3). All studies were non-randomised comparisons in adults. Three studies compared radioactive iodine with thyroidectomy and the remaining five studies compared people treated with radioactive iodine, with the general population.

Where there were studies assessing the same cohort these were handled so as to minimise double counting within the same meta-analysis.

Death from cancer was considered a surrogate outcome for cancer diagnoses and was extracted if diagnoses of cancer were not available for that cohort comparison. This outcome was downgraded for indirectness.

2.4.2 Excluded studies

See the excluded studies list in Appendix J:.

2.4.3 Summary of clinical studies included in the evidence review

Table 10: Summary of studies included in the evidence review

Study	Intervention and comparison	Population	Outcomes	Comments
Franklyn 1999[38]	Radioactive iodine, n = 7417 Fixed dose, mean 308Mbq (SD 232), 84.9% received only one dose Age, sex and year matched SIR	People with hyperthyroidism, treated in West Midlands with radioactive iodine (mean age at treatment 56.6) Cohort treated between 1950-1991 UK	Overall cancer incidence Site specific cancer incidence Mean follow-up 9.7 years,	
Franklyn 2005[40]	Radioactive iodine, n = 2668 Fixed dose either 185 or 370MBq, 84.3% received one dose only Age, sex and year matched SMRs	People with hyperthyroidism, treated in West Midlands with radioactive iodine (median age at treatment start 62) Cohort treated between 1984-2002 UK	Cancer mortality Median follow-up 5.6 years	542 person overlap with Franklyn 1999, outcome different
Giesecke 2018[42]	Radioactive iodine, n = 10250 Dose not specified Thyroidectomy, n = 742 Surgery not specified	People with hyperthyroidism (mean age of RAI group 64, mean of surgery group 47) Cohort treated between 1976-2013 Sweden	Cancer mortality Mean follow-up 16.3 years	Adjusted for potential confounders in regression of age at treatment, gender, year of treatment, aetiology, co-existing conditions
Goldman 1988[44]	Radioactive iodine, n = 1762 Dose not specified	Women with hyperthyroidism (age not stated) treated with	Overall cancer incidence Site specific cancer	

Study	Intervention and comparison	Population	Outcomes	Comments
	Age, sex, race, year matched SIRs for Connecticut	RAI at Mass. Gen. Hospital Cohort treated between 1946 and 1964 USA	incidence Mean follow-up 17.2 years	
Hoffman 1982[59]	Radioactive iodine, n = 1005 Mean dose 10.6mCi (~392MBq), mean number of doses 1.2 Thyroidectomy, n = 2141 Surgery not specified	White women with hyperthyroidism treated by Mayo clinic (mean age of RAI group at Tx 56.8, surgery 45.7) Cohort treated between 1946 and 1964 USA	Overall cancer incidence Site specific cancer incidence Mean follow-up 15 years for RAI group, 21 years for surgical group	Adjusted for age, year of treatment, duration of follow-up
Holm 1991[62]	Radioactive iodine, n = 10207 Mean dose 506MBq Age, sex, region, year matched incidence for whole of Sweden	People with hyperthyroidism (mean age 57, range 13-74) Cohort treated between 1950 and 1975 Sweden	Overall cancer incidence Site specific cancer incidence Mean follow-up 15 years	
Metso 2007[90]	Radioactive iodine, n = 2793 Mean dose 305MBq Age, sex matched control from Finnish population register	People with hyperthyroidism treated with RAI at Tampere hospital (median age 62 years) Cohort treated between 1965 and 2002 Finland	Overall cancer incidence Site specific cancer incidence Mean follow-up 9.8 years for patients and 10.0 years for controls	
Ryodi 2015[115]	Radioactive iodine, n = 1814	People with hyperthyroidism	All cancer diagnoses	Unspecified overlap with Metso

Study	Intervention and comparison	Population	Outcomes	Comments
	Dose not specified Thyroidectomy, n = 4334 Surgery not specified	(median age of radioactive iodine group 59, median age of thyroidectomy group 46) Cohort treated between 1986-2007 Finland	Median follow-up 10 years	2007, however comparison different Adjusted for aetiology, age, and gender

See Appendix D: for full evidence tables.

2.4.4 Quality assessment of clinical studies included in the evidence review

Table 11: Clinical evidence summary: radioactive iodine vs surgery

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Surgery	Risk difference with Radioactive iodine (95% CI)
Total cancer diagnoses (RR)	3146 (1 study) 15 years	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	RR 1.00 (0.7 to 1.43)	115 per 1000	0 fewer per 1000 (from 35 fewer to 49 more)
Total cancer diagnoses (HR)	6148 (1 study) 10 years	⨁⊖⊖⊖ VERY LOW[1] due to risk of bias	HR 1.03 (0.86 to 1.23)	-[3]	Not estimable
Total cancer mortality	10992 (1 study) 16.3 years	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	HR 0.96 (0.73 to 1.26)	-[3]	Not estimable
Lip, oral, pharynx cancer diagnoses	3146 (1 study) 15 years	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, imprecision	RR 1.3 (0.2 to 8.45)	4 per 1000	1 more per 1000 (from 3 fewer to 28 more)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Surgery	Risk difference with Radioactive iodine (95% CI)
	(1 study) 15 years	VERY LOW1,2 due to risk of bias, imprecision	(0.6 to 2.02)	24 per 1000	2 more per 1000 (from 10 fewer to 24 more)
Respiratory cancer diagnoses	3146 (1 study) 15 years	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 1.3 (0.4 to 4.23)	7 per 1000	2 more per 1000 (from 4 fewer to 23 more)
Breast cancer diagnoses	3146 (1 study) 15 years	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 0.8 (0.5 to 1.28)	34 per 1000	7 fewer per 1000 (from 17 fewer to 10 more)
Genital cancer diagnoses	3146 (1 study) 15 years	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 1.1 (0.4 to 3.02)	21 per 1000	2 more per 1000 (from 13 fewer to 42 more)
Kidney and bladder cancer diagnoses	3146 (1 study) 15 years	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 3.4 (0.5 to 23.12)	2 per 1000	5 more per 1000 (from 1 fewer to 42 more)
Melanoma diagnoses	3146 (1 study) 15 years	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 0 (0 to 7.8)	0 per 1000	1 fewer per 1000 (from 1 fewer to 3 more)
CNS cancer diagnoses	3146 (1 study) 15 years	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 0.3 (0.05 to 1.9)	3 per 1000	2 fewer per 1000 (from 3 fewer to 3 more)
Thyroid cancer diagnoses	3146 (1 study) 15 years	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias,	RR 9.1 (1.2 to 69.01)	0 per 1000	4 more per 1000 (from 0 more to 34 more)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Surgery	Risk difference with Radioactive iodine (95% CI)
		imprecision			
Other solid tumour diagnoses	3146 (1 study) 15 years	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 0.3 (0.02 to 4.3)	3 per 1000	2 fewer per 1000 (from 3 fewer to 9 more)
Lymphatic cancer diagnoses	3146 (1 study) 15 years	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 0.3 (0.02 to 3.7)	3 per 1000	2 fewer per 1000 (from 3 fewer to 9 more)
Leukaemia diagnoses	3146 (1 study) 15 years	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 0.6 (0.16 to 2.2)	5 per 1000	2 fewer per 1000 (from 4 fewer to 6 more)

1 Default starting quality of low overall due to selection bias in non-randomised studies. Downgraded further for risk of bias if the majority of evidence was at additional risk of bias, either once if high risk of bias or twice if very high risk of bias
2 Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs
3 No control group risk provided

Table 12: Clinical evidence summary: radioactive iodine treated population vs general population

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with General population	Risk difference with Radioactive iodine (95% CI)
Total cancer diagnoses	26485 (5 studies) 5-17 years	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, inconsistency	Rate ratio 0.99 (0.83 to 1.18)	74 per 1000	1 fewer per 1000 (from 13 fewer to 13 more)
Lip, oral, pharynx cancer diagnoses	23210 (3 studies) 5-15 years	⨁⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision	Rate ratio 0.92 (0.57 to	1 per 1000	0 fewer per 1000 (from 0 fewer to 0 more)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with General population	Risk difference with Radioactive iodine (95% CI)
			1.49)		
Salivary gland cancer diagnoses	15793 (2 studies) 10-15 years	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision	Rate ratio 1.88 (0.33 to 10.62)	0 per 1000	0 more per 1000 (from 0 fewer to 1 more)
Digestive organs and peritoneum cancer diagnoses	23817 (4 studies) 5-17 years	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, inconsistency	Rate ratio 1.06 (0.87 to 1.30)	27 per 1000	2 more per 1000 (from 4 fewer to 8 more)
Bone, connective tissue and skin cancer diagnoses	13003 (2 studies) 5-10 years	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	Rate ratio 0.88 (0.69 to 1.14)	13 per 1000	2 fewer per 1000 (from 4 fewer to 2 more)
Breast cancer diagnoses	23817 (4 studies) 5-17 years	⊕⊖⊖⊖ VERY LOW1 due to risk of bias	Rate ratio 1.09 (0.97 to 1.22)	17 per 1000	2 more per 1000 (from 1 fewer to 4 more)
Brain and other CNS cancer diagnoses	23817 (4 studies) 5-17 years	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision	Rate ratio 1.46 (1.03 to 2.06)	3 per 1000	1 more per 1000 (from 0 more to 3 more)
Respiratory cancer diagnoses	23210 (3 studies) 5-17 years	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, inconsistency, imprecision	Rate ratio 0.84 (0.52 to 1.35)	9 per 1000	1 fewer per 1000 (from 4 fewer to 3 more)
Genitourinary cancer diagnoses	23210 (3 studies) 5-17 years	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, inconsistency, imprecision	Rate ratio 0.95 (0.73 to 1.24)	16 per 1000	1 fewer per 1000 (from 4 fewer to 4 more)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with General population	Risk difference with Radioactive iodine (95% CI)
	(3 studies) 5-17 years	VERY LOW1 due to risk of bias	2.17 (1.36 to 3.48)	1 per 1000	1 more per 1000 (from 0 more to 2 more)
Haematopoietic cancer diagnoses	23210 (3 studies) 5-17 years	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, inconsistency, imprecision	Rate ratio 0.81 (0.56 to 1. 19)	5 per 1000	1 fewer per 1000 (from 2 fewer to 1 more)
Kidney cancer diagnoses	15793 (2 studies) 10-15 years	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision	Rate ratio 1.62 (1.18 to 2.24)	4 per 1000	2 more per 1000 (from 1 more to 5 more)
Parathyroid cancer diagnoses	10207 (1 study) 15 years	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision	Rate ratio 1.6 (0.9 to 2.84)	2 per 1000	1 more per 1000 (from 0 fewer to 4 more)
Prostate cancer diagnoses	5586 (1 study) 10 years	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision	Rate ratio 1.3 (0.69 to 2.45)	37 per 1000	11 more per 1000 (from 11 fewer to 54 more)

1 Default starting quality of low overall due to selection bias in non-randomised studies. Downgraded further for risk of bias if the majority of evidence was at additional risk of bias, either once if high risk of bias or twice if 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

See Appendix F: for full GRADE tables.

2.5 Economic evidence

The committee agreed that health economic studies would not be relevant to this review question, and so were not sought.

2.6 Evidence statements

2.6.1 Clinical evidence statements

2.6.1.1 Radioactive iodine vs surgery

No clinically important difference was identified for total cancer diagnoses (2 studies, very low quality), total cancer mortality (1 study, very low quality), lip/oral/pharynx cancer diagnoses (1 study, very low quality), digestive organ and peritoneum cancer diagnoses (1 study, very low quality), respiratory cancer diagnoses (1 study, very low quality), breast cancer diagnoses (1 study, very low quality), genital cancer diagnoses (1 study, very low quality), kidney and bladder cancer diagnoses (1 study, very low quality), melanoma diagnoses (1 study, very low quality), CNS cancer diagnoses (1 study, very low quality), thyroid cancer diagnoses (1 study, very low quality), other solid tumour diagnoses (1 study, very low quality), lymphatic cancer diagnoses (1 study, very low quality), leukaemia diagnoses (1 study, very low quality).

2.6.1.2 Radioactive iodine vs general population

No clinically important difference was identified for total cancer diagnoses (5 studies, very low quality), lip/oral/pharynx cancer diagnoses (3 studies, very low quality), salivary gland cancer diagnoses (2 studies, very low quality), digestive organ and peritoneum cancer diagnoses (4 studies, very low quality), bone/connective tissue/skin cancer diagnoses (2 studies, very low quality), breast cancer diagnoses (4 studies, very low quality), brain and other CNS cancer diagnoses (4 studies, very low quality), respiratory cancer diagnoses (3 studies, very low quality), genitourinary cancer diagnoses (3 studies, very low quality), thyroid cancer diagnoses (1 study, very low quality), haematopoietic cancer diagnoses (3 studies, very low quality), kidney cancer diagnoses (2 studies, very low quality), parathyroid cancer diagnoses (1 study, very low quality).

There was a clinically important harm of radioactive iodine for prostate cancer diagnoses (1 study, very low quality).

2.7 The committee's discussion of the evidence

2.7.1 Interpreting the evidence

2.7.1.1 The outcomes that matter most

Drugs vs Surgery vs Radioactive Iodine

The committee agreed that the critical outcomes for this review were mortality and quality of life. Important outcomes for all interventions included thyroid ophthalmopathy, euthyroidism, hypothyroidism, relapse of hyperthyroidism, cardiovascular morbidity, arrhythmia, osteoporosis, cognitive impairment, pain, symptom scores, experience of care, healthcare contacts. Important intervention specific outcomes were recurrent laryngeal nerve damage, hypocalcaemia, hypoparathyroidism, bleeding, infection, agranulocytosis, liver failure, minor

drug related adverse effects, teratogenesis, infertility, malignancy, thyrotoxic storm, growth abnormalities.

Radioactive Iodine safety

The committee considered cancer diagnoses and infertility to be critical outcomes for this review. No evidence was found for infertility.

2.7.1.2 The quality of the evidence

Drugs vs Surgery vs Radioactive Iodine

The quality of the evidence in this review ranged from high to very low quality, with the majority being moderate or low quality. Evidence was typically downgraded for imprecision as studies were often small, some comparisons were downgraded for inconsistency that could not be explained by any protocol subgroup analyses and some comparisons were downgraded for risk of bias. The trials included in this review generally had long follow-up periods, with some participants being followed for up to 21 years.

Thyroid status and ophthalmopathy at end of follow-up were the most commonly reported outcome. There was no quality of life evidence identified. The majority of outcomes included in the protocol by the committee were not reported.

The comparison between radioactive iodine and antithyroid drugs, used for first-line treatment of Graves' disease, included the most evidence. Comparisons involving surgery or second line treatment were only supported by one study each.

No evidence was identified in children or older adults. No evidence was identified in studies explicitly in toxic multinodular goitre and there was only one small study in people who had failed first line treatment.

The committee noted that the studies included in this review were not designed to capture the rare but well established adverse events of some treatment options (for example agranulocytosis with antithyroid drugs).

The committee noted that the doses of radioactive iodine used in most studies were lower than what would be used in the UK currently. Qualitatively, higher doses would be expected to lead to more hypothyroidism and euthyroidism and less hyperthyroidism. Higher doses could also lead to more adverse events, although these were not identified in this review.

Radioactive Iodine safety

The majority of the evidence was very low quality due to the non-randomised nature of the included studies. Beyond the lack of randomisation, studies that compared radioactive iodine with surgery were more informative as they reduced the confounding effect of the underlying thyroid disease (as opposed to the studies that compared cancer diagnoses between a radioactive iodine treated group and the general population). The majority of studies included a population who had been treated many years ago, some cohorts including participants treated as far back as 1946. The doses and strategies of radioactive iodine (for example whether a fixed administered activity or calculated absorbed dose was used) were not always provided but generally appeared to be a fixed approach and using lower doses (for example ~300MBq) than those used in the UK currently (typically 400-600MBq).

The committee agreed that there were limitations to the evidence available but also noted that the studies were large, had long follow-up times and it is unlikely that RCTs that are as large and lengthy in follow-up will ever be conducted. Nevertheless, they agreed that a registry of patients receiving RAI would further develop our understanding of the risks and benefits associated with RAI therapy and decided to make a research recommendation.

All the included evidence was on adults, there was no evidence to consider in children.

2.7.1.3 Benefits and harms

Drugs vs surgery vs radioactive Iodine

The committee noted that antithyroid drugs could be used in two main ways, either to control thyrotoxicosis prior to treatment with radioactive iodine or surgery or as definitive treatment. The evidence identified in the review assessed the efficacy of the latter which is the focus of the discussion below. However the committee agreed based on their experience that the use of antithyroid drugs to control thyrotoxicosis in the acute period is important to optimise later treatment, prevent acute illness if thyrotoxicosis is severe and in some circumstances to address delays in access to other definitive treatments.

The evidence shows that radioactive iodine has a clinically important harm compared with antithyroid drugs as a definitive treatment for ophthalmopathy but a clinically important benefit in terms of reducing persistence or recurrence of hyperthyroidism.

Compared with antithyroid drugs, radioactive iodine also appeared to lead to more people ending up in a hypothyroid state as opposed to euthyroid. The committee discussed the outcomes of hypothyroidism and euthyroidism. Euthyroidism is seen as a preferential goal of treatment by some people with thyroid disease, and eliminates the need for concurrent thyroid function replacement with thyroid hormone replacement. However the committee was aware of some evidence (not the focus of this review) that long term outcomes for people who achieve hypothyroidism after radioactive iodine are better than for those who are euthyroid. The committee noted that current guidance by other groups is to aim for hypothyroidism when using radioactive iodine. Committee members in primary care noted that from their experience, the people they treated for hyperthyroidism with antithyroid drugs as definitive treatment were generally more satisfied with their care than the people they treated for hypothyroidism (secondary to radioactive iodine or surgery for hyperthyroidism).

There was less evidence available comparing surgery to either modality. In general surgery appeared to have a clinically important benefit over radioactive iodine or antithyroid drugs in terms of the likelihood of relapse or persistence of hyperthyroidism; however the smaller trials made this difficult to interpret. The committee noted that although there was no evidence in this review on hypothyroidism as a result of surgery, this is conceptually a likely outcome. The one study in this review reporting on surgical outcomes assessed the efficacy of subtotal thyroidectomy, as opposed to total thyroidectomy as in the economic evidence. As discussed in the review of different types of surgery, these two options are likely to have different benefits and harms.

Beyond the impact of ophthalmopathy and thyroid state, the review did not identify definitive evidence on the harms of radioactive iodine, antithyroid drugs or surgery. The committee agreed that each form of treatment is associated with some harm. Some of these harms are more definitive than others. Surgery is associated with the general harms of surgery (for example bleeding, infection) as well as specific harms related to surgery on the thyroid gland (for example hypoparathyroidism and recurrent laryngeal nerve damage). Antithyroid drugs have a combination of common minor adverse events (for example skin rash) and rare but severe adverse events (for example agranulocytosis and liver failure) which are documented in the summary of product characteristics. Radioactive iodine treatment has theoretical harms beyond those identified in this review, in terms of secondary malignancies and effects on fertility or teratogenesis. None of these harms were identified in the RCTs in this review and the committee's view overall was that while these were important risks to discuss with people considering treatment, there was not information available on their likelihood.

The committee noted, based on their experience, that there may be particular features of a person's hyperthyroidism that may suggest one treatment option is preferable to others. If

there was any uncertainty around the potential for thyroid cancer or if there were significant compressive symptoms from a large goitre, then surgery was typically considered the most appropriate option. If there was a significant degree of pre-existing ophthalmopathy this may promote treatment options other than radioactive iodine. If people's hyperthyroidism generally appeared likely to respond well to antithyroid drug treatment, this may make them a better candidate for first-line definitive treatment with drugs as opposed to potentially causing long term hypothyroidism with either radioactive iodine or surgery. The committee noted that there was no evidence in this review to suggest which groups might respond particularly well to antithyroid drugs. In their experience, people with very mild hyperthyroidism and in particular T3 hyperthyroidism did tend to respond well to antithyroid drugs.

The committee discussed the extrapolation of evidence and experience from adults to children. Concerns over potential adverse effects of definitive treatment with radioactive iodine or surgery were generally greater for children than adults. Surgery may be technically more demanding in children. The potential long term risks of radioactive iodine in terms of secondary malignancy are more relevant in children, given their greater life expectancy after treatment compared with older adults. However at the same time, children and their families are often keen to explore definitive treatment options. From the committee's experience, hyperthyroidism in children may be more aggressive than in adults and require lengthier treatment with antithyroid drugs (up to 10 years in children as opposed to 12-18 months in adults).

Radioactive iodine safety

Overall the evidence in this review did not show a clinically important harm of radioactive iodine treatment compared with either surgery or a general population in terms of increased risk of cancer diagnoses. There was no clinically important effect for overall cancer diagnoses, the outcome with the greatest event rates in both arms. When considering site specific cancer diagnoses, due to the much smaller event rates there was generally more imprecision and lower quality evidence with relative effects more likely to appear to show an effect but the absolute effects remained small, with all but one remaining below the threshold of 10 per 1000 people treated. The committee agreed that the one outcome, for which this threshold was breached, prostate cancer diagnoses in the radioactive iodine versus general population comparison, was likely to reflect statistical uncertainty more than a true effect and noted the very low quality of the evidence.

The committee agreed that there was insufficient evidence to determine in this review if dosing strategy affected safety as the studies generally did not provide adequate information on the radioactive iodine strategies used.

Balanced against this evidence of no important harm of radioactive iodine, the committee noted the underlying biological principles that any exposure to radiation is likely to increase cancer risk to some degree. However the evidence in this review suggests that the risk associated with the radiation involved in treatment of thyroid disease is not clinically impactful.

2.7.2 Cost effectiveness and resource use

Resource use implications were considered through the published cost-effectiveness evidence included in the review.

This was a UK cost-utility analysis that compared three options for the management of thyrotoxicosis secondary to Graves' disease: radioactive iodine (RAI); antithyroid drugs (ATD); and surgery (total thyroidectomy).

The analysis found that RAI was the most cost effective option at a cost effectiveness threshold of £20,000 per QALY gained. RAI had the lowest mean cost per patient over a lifetime horizon (£5,425) and a mean 34.73 QALYs per patient. Total thyroidectomy had

higher costs (£7,115) and lower QALYs (33.93 QALYs) than RAI. ATDs had higher costs (£16,866) than RAI but also higher QALYs (35.17 QALYs); however, it had an incremental cost effectiveness ratio compared to RAI of £26,279 per QALY gained and so was not considered cost effective.

The committee noted that the results of the economic evidence were in line with the clinical evidence. This supported a strong recommendation to offer radioactive iodine as the first line treatment option for the management of thyrotoxicosis secondary to Graves' disease, unless it is unsuitable (for example if there are concerns about compression, malignancy is suspected or if the patient is pregnant or trying to become pregnant or father a child) or if antithyroid drugs are likely to achieve remission. The committee noted that if the latter was likely to be the case, a choice of antithyroid drugs and radioactive iodine should be offered as first-line treatment. For example, people with mild and uncomplicated Graves' disease whom are likely to achieve remission with a course of antithyroid drugs are unlikely to be rendered hypothyroid, reducing the need for long-term hormone replacement therapy which in turn saves money and improves patients' quality of life.

The committee noted that the model accurately captures the following key adverse events associated with each of the three interventions: hypothyroidism secondary to total thyroidectomy, the excess risk of ophthalmopathy when using radioactive iodine and increased risk of relapse when using antithyroid drugs. However, it was noted that some potential adverse events have not been reflected in the model structure, e.g. malignancy, thyroid eye disease. This was justified in the study, though, as the authors explained that the evidence supporting causal association between the use of radioactive iodine and malignancy is limited which was confirmed by the committee during their discussions. However, as a result of this concern, the committee chose to restrict the recommendation to people in whom there is no risk of malignancy, thyroid eye disease or compression. The risk of infertility was another potential adverse event, not captured in the model, and relating to the use of radioactive iodine in women of childbearing age. Therefore, the committee agreed that it was important to discuss treatment options with people with Graves' disease to minimise these risks.

There was no economic or clinical evidence for the management of thyrotoxicosis in people with toxic nodular goitre. Hence, the committee extrapolated the findings from people with Graves' disease and made a recommendation to offer radioactive iodine as first line treatment except in instances where there are concerns around malignancy.

In children, the committee were uncertain about the long-term health risk associated with radioactive iodine and surgery, and agreed to offer antithyroid drugs as first line treatment, which is in line with current practice and unlikely to have a substantial cost impact. However, the committee noted that definitive options should be discussed with a multi-disciplinary team especially when they have relapsed hyperthyroidism after a course of antithyroid drugs or in children with a single toxic nodule. The population of children with single toxic nodule is very small hence unlikely to result in a cost impact.

Overall, the recommendation for the use of RAI as first line is a change to current practice, which is likely to be cost effective as shown by the economic evidence and agreed by the committee. Furthermore, in children due to the uncertainty around the potential risk and benefit around radioactive iodine treatment, the cost-effectiveness was considered uncertain.

2.7.3 Other factors the committee took into account

The committee noted that none of the currently available treatment options addressed the potential underlying causes of hyperthyroidism (for example the immunological basis for Graves' disease). While immunomodulatory treatment options were not a focus of this review

and therefore specific research recommendations could not be made, the committee were keen to see this area be developed in the future.

The committee noted that although pregnancy is outside the scope of this guideline, radioactive iodine would not be considered appropriate for anyone considering pregnancy, currently pregnant or breast-feeding.

The committee made recommendations on toxic multinodular goitre based on extrapolations from the evidence on Graves' disease (noting that some studies in predominantly Graves' disease populations did include a minority of people with toxic multinodular goitre) and on their own experience. The committee's experience was that in most cases antithyroid drugs would not be an appropriate option for this population, however, if radioactive iodine or surgery are not suitable then antithyroid drugs are likely to be needed.

The committee noted that by the point in the treatment pathway that people arrive at radioactive iodine currently, they have typically been started on antithyroid drugs in primary care. For example in some places people may be prescribed antithyroid drugs in primary care as a stop gap measure until specialist referral is available. However in other places primary healthcare professionals are unwilling to initiate antithyroid drugs without specialist input. The committee agreed that it was unacceptable to leave people with thyrotoxicosis without antithyroid treatment and agreed that antithyroid drugs along with supportive treatment should be considered for adults with hyperthyroidism who are waiting for specialist assessment and further treatment.The review on the use of radioactive iodine considers this issue further.

References

1. Abraham-Nordling M, Törring O, Hamberger B, Lundell G, Tallstedt L, Calissendorff J et al. Graves' disease: a long-term quality-of-life follow up of patients randomized to treatment with antithyroid drugs, radioiodine, or surgery. Thyroid. 2005; 15(11):1279-1286

2. 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

3. Abraham-Nordling M, Wallin G, Träisk F, Berg G, Calissendorff J, Hallengren B et al. Thyroid-associated ophthalmopathy; quality of life follow-up of patients randomized to treatment with antithyroid drugs or radioiodine. European Journal of Endocrinology. 2010; 163(4):651-657

4. 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.

5. 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

6. 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

7. 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

8. 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

9. Angusti T, Codegone A, Pellerito R, Favero A. Thyroid cancer prevalence after radioiodine treatment of hyperthyroidism. Journal of Nuclear Medicine. 2000; 41(6):1006-9

10. 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

11. Azizi F, Ataie L, Hedayati M, Mehrabi Y, Sheikholeslami F. Effect of long-term continuous methimazole treatment of hyperthyroidism: comparison with radioiodine. European Journal of Endocrinology. 2005; 152(5):695-701

12. 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

13. 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

14. 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

15. 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

16. Bartalena L, Marcocci C, Bogazzi F, Manetti L, Tanda ML, Dell'Unto E et al. Relation between therapy for hyperthyroidism and the course of Graves' ophthalmopathy. New England Journal of Medicine. 1998; 338(2):73-78

17. 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

18. 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

19. 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

20. 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

21. 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

22. 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

23. 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

24. 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

25. 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

26. Cevallos JL, Hagen GA, Maloof F, Chapman EM. Low-dosage 131-I therapy of thyrotoxicosis (diffuse goiters). A five-year follow-up study. New England Journal of Medicine. 1974; 290(3):141-3

27. Chao M, Jiawei X, Guoming W, Jianbin L, Wanxia L, Driedger A et al. Radioiodine treatment for pediatric hyperthyroid Grave's disease. European Journal of Pediatrics. 2009; 168(10):1165-1169

28. Chen DY, Jing J, Schneider PF, Chen TH. Comparison of the long-term efficacy of low dose ^{131}I versus antithyroid drugs in the treatment of hyperthyroidism. Nuclear Medicine Communications. 2009; 30(2):160-168

29. 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

30. 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

31. 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

32. 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

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

34. Department of Health. NHS reference costs 2016-17. 2017. Available from: https://www.gov.uk/government/collections/nhs-reference-costs Last accessed: 15/02/2019

35. Donovan PJ, McLeod DS, Little R, Gordon L. Cost-utility analysis comparing radioactive iodine, anti-thyroid drugs and total thyroidectomy for primary treatment of Graves' disease. European Journal of Endocrinology. 2016; 175(6):595-603

36. 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

37. 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

38. Franklyn JA, Maisonneuve P, Sheppard M, Betteridge J, Boyle P. Cancer incidence and mortality after radioiodine treatment for hyperthyroidism: a population-based cohort study. Lancet. 1999; 353(9170):2111-5

39. Franklyn JA, Maisonneuve P, Sheppard MC, Betteridge J, Boyle P. Mortality after the treatment of hyperthyroidism with radioactive iodine. New England Journal of Medicine. 1998; 338(11):712-8

40. Franklyn JA, Sheppard MC, Maisonneuve P. Thyroid function and mortality in patients treated for hyperthyroidism. JAMA. 2005; 294(1):71-80

41. 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

42. Giesecke P, Frykman V, Wallin G, Lonn S, Discacciati A, Torring O et al. All-cause and cardiovascular mortality risk after surgery versus radioiodine treatment for hyperthyroidism. British Journal of Surgery. 2018; 105(3):279-286

43. 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

44. Goldman MB, Maloof F, Monson RR, Aschengrau A, Cooper DS, Ridgway EC. Radioactive iodine therapy and breast cancer. A follow-up study of hyperthyroid women. American Journal of Epidemiology. 1988; 127(5):969-80

45. 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

46. 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

47. Hall P, Berg G, Bjelkengren G, Boice JD, Jr., Ericsson UB, Hallquist A et al. Cancer mortality after iodine-131 therapy for hyperthyroidism. International Journal of Cancer. 1992; 50(6):886-90

48. Hall P, Boice JD, Jr., Berg G, Bjelkengren G, Ericsson UB, Hallquist A et al. Leukaemia incidence after iodine-131 exposure. Lancet. 1992; 340(8810):1-4

49. Hall P, Holm LE. Cancer in iodine-131 exposed patients. Journal of Endocrinological Investigation. 1995; 18(2):147-9

50. Hall P, Holm LE. Late consequences of radioiodine for diagnosis and therapy in Sweden. Thyroid. 1997; 7(2):205-8

51. Hall P, Lundell G, Holm LE. Mortality in patients treated for hyperthyroidism with iodine-131. Acta Endocrinologica. 1993; 128(3):230-4

52. 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

53. 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

54. 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

55. Hieu TT, Russell AW, Cuneo R, Clark J, Kron T, Hall P et al. Cancer risk after medical exposure to radioactive iodine in benign thyroid diseases: A meta-analysis. Endocrine-Related Cancer. 2012; 19(5):645-655

56. 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

57. Hoffman DA, McConahey WM. Breast cancer following iodine-131 therapy for hyperthyroidism. Journal of the National Cancer Institute. 1983; 70(1):63-7

58. Hoffman DA, McConahey WM, Diamond EL, Kurland LT. Mortality in women treated for hyperthyroidism. American Journal of Epidemiology. 1982; 115(2):243-54

59. Hoffman DA, McConahey WM, Fraumeni JF, Jr., Kurland LT. Cancer incidence following treatment of hyperthyroidism. International Journal of Epidemiology. 1982; 11(3):218-24

60. Holm LE. Thyroid cancer after exposure to radioactive 131I. Acta Oncologica. 2006; 45(8):1037-40

61. Holm LE, Dahlqvist I, Israelsson A, Lundell G. Malignant thyroid tumors after iodine-131 therapy: a retrospective cohort study. New England Journal of Medicine. 1980; 303(4):188-91

62. Holm LE, Hall P, Wiklund K, Lundell G, Berg G, Bjelkengren G et al. Cancer risk after iodine-131 therapy for hyperthyroidism. Journal of the National Cancer Institute. 1991; 83(15):1072-7

63. 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

64. 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

65. 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

66. 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

67. 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

68. Journy NMY, Bernier MO, Doody MM, Alexander BH, Linet MS, Kitahara CM. Hyperthyroidism, hypothyroidism, and cause-specific mortality in a large cohort of women. Thyroid. 2017; 27(8):1001-1010

69. 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

70. Kansara S, Kotwal N, Kumar K, Singh Y, Upreti V, Nachankar A. Effect of Antithyroid Therapies on Bone and Body Composition: A Prospective, Randomized, Clinical Study Comparing Antithyroid Drugs with Radioiodine Therapy. Indian Journal of Endocrinology and Metabolism. 2017; 21(4):531-534

71. 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

72. 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

73. 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

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

75. 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

76. 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

77. 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.

78. 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

79. 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

80. 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.

81. 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: https://dx.doi.org/10.1002/14651858.CD010094.pub2.

82. 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

83. 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

84. 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

85. 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

86. 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

87. McTiernan AM, Weiss NS, Daling JR. Incidence of thyroid cancer in women in relation to previous exposure to radiation therapy and history of thyroid disease. Journal of the National Cancer Institute. 1984; 73(3):575-81

88. 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

89. Metso S, Auvinen A, Huhtala H, Salmi J, Oksala H, Jaatinen P. Erratum: Increased cancer incidence after radioiodine treatment for hyperthyroidism. Cancer. 2007; 110(8):1875

90. Metso S, Auvinen A, Huhtala H, Salmi J, Oksala H, Jaatinen P. Increased cancer incidence after radioiodine treatment for hyperthyroidism. Cancer. 2007; 109(10):1972-9

91. Metso S, Jaatinen P, Huhtala H, Auvinen A, Oksala H, Salmi J. Erratum: Increased cardiovascular and cancer mortality after radioiodine treatment for hyperthyroidism. Journal of Clinical Endocrinology and Metabolism. 2007; 92(10):4008

92. Metso S, Jaatinen P, Huhtala H, Auvinen A, Oksala H, Salmi J. Increased cardiovascular and cancer mortality after radioiodine treatment for hyperthyroidism. Journal of Clinical Endocrinology and Metabolism. 2007; 92(6):2190-6

93. Metso S, Jaatinen P, Huhtala H, Luukkaala T, Oksala H, Salmi J. Long-term follow-up study of radioiodine treatment of hyperthyroidism. Clinical Endocrinology. 2004; 61(5):641-648

94. 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)

95. 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

96. 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

97. 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

98. 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

99. 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

100. 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

101. Patel NN, Abraham P, Buscombe J, Vanderpump MP. The cost effectiveness of treatment modalities for thyrotoxicosis in a U.K. center. Thyroid. 2006; 16(6):593-8

102. 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

103. 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

104. 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

105. 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

106. 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

107. 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

108. 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

109. 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

110. 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

111. 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

112. 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

113. 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

114. Ron E, Doody MM, Becker DV, Brill AB, Curtis RE, Goldman MB et al. Cancer mortality following treatment for adult hyperthyroidism. Journal of the American Medical Association. 1998; 280(4):347-355

115. Ryodi E, Metso S, Jaatinen P, Huhtala H, Saaristo R, Valimaki M et al. Cancer Incidence and Mortality in Patients Treated Either With RAI or Thyroidectomy for Hyperthyroidism. Journal of Clinical Endocrinology and Metabolism. 2015; 100(10):3710-7

116. 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

117. 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

118. 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

119. 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

120. Singer RB. Long-term comparative cancer mortality after use of radio-iodine in the treatment of hyperthyroidism, a fully reported multicenter study. Journal of Insurance Medicine (Seattle). 2001; 33(2):138-42

121. Singer RB. Long-term comparative mortality in hyperthyroid patients treated with radio-iodine, a cohort study in England. Journal of Insurance Medicine (Seattle). 2001; 33(2):133-7

122. 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.

123. 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

124. Tallstedt L, Lundell G, Tørring O, Wallin G, Ljunggren JG, Blomgren H et al. Occurrence of ophthalmopathy after treatment for Graves' hyperthyroidism. The Thyroid Study Group. New England Journal of Medicine. 1992; 326(26):1733-1738

125. 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

126. 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

127. Törring O, Tallstedt L, Wallin G, Lundell G, Ljunggren JG, Taube A et al. Graves' hyperthyroidism: treatment with antithyroid drugs, surgery, or radioiodine--a prospective, randomized study. Thyroid Study Group. Journal of Clinical Endocrinology and Metabolism. 1996; 81(8):2986-2993

128. Träisk F, Tallstedt L, Abraham-Nordling M, Andersson T, Berg G, Calissendorff J et al. Thyroid-associated ophthalmopathy after treatment for Graves' hyperthyroidism with antithyroid drugs or iodine-131. Journal of Clinical Endocrinology and Metabolism. 2009; 94(10):3700-3707

129. 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

130. Verburg FA, Luster M, Lassmann M, Reiners C. 131-I therapy in patients with benign thyroid disease does not conclusively lead to a higher risk of subsequent malignancies. Nuklearmedizin. 2011; 50(3):93-9; quiz N20

131. 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

132. 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

133. 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

134. 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

135. 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

136. Yuan J, Lu X, Yue Y. Comparison of curative effect of 131I and antithyroid drugs in Graves' disease: a meta analysis. Minerva Endocrinologica. 2017; Epublication

Appendices

Appendix A: Review protocols

Table 13: Review protocol: Management of Thyrotoxicosis: Drugs vs Surgery vs Radioactive iodine

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)	• Comparisons between modalities • Comparisons between submodalities

	/ control or 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 (radioactive iodine)** • 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 sensitivity / subgroup analysis, or meta-regression	**Stratifications** • 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 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 appendix D 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 14: Review protocol: Radioactive iodine safety

ID	Field	Content
I	Review question	What are the long term adverse events of radioactive iodine treatment for thyrotoxicosis?
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	To determine the long term adverse event profile of radioactive iodine treatment for thyrotoxicosis
IV	Eligibility criteria – population / disease / condition / issue / domain	• No population restrictions (see below for prioritising of evidence)
V	Eligibility criteria – intervention(s) / exposure(s) / prognostic factor(s)	• Radioactive iodine
VI	Eligibility criteria – comparator(s) / control or reference (gold) standard	• Antithyroid drug treatment • Surgical treatment • Healthy controls (see below for prioritising of evidence)
VII	Outcomes and prioritisation	• Cancer ○ Overall ○ Organs group specific • Infertility
VIII	Eligibility criteria – study design	• Evidence will be considered according to the following hierarchy: ○ Cohort studies with hyperthyroid controls and adequate adjustment for key confounders (age, smoking) ○ Cohort studies with hyperthyroid controls and without adequate adjustment for key confounders ○ Cohort studies with healthy controls and adequate adjustment for key confounders (age, smoking) ○ Cohort studies with healthy controls without adequate adjustment for key confounders
IX	Other inclusion exclusion criteria	• Only included if: ○ For adults sample size >1000 ○ Length of follow-up >5 years
X	Proposed sensitivity / subgroup analysis, or meta-regression	**Stratifications** • Age – infants (<4), children and young people (4-18), adults (>18-65), older adults (>65) **Subgroup analyses** • Dose of radioactive iodine – fixed administered activity 200-<400 MBq, fixed 400-800 MBq, calculated absorbed dose strategy
XI	Selection process –	• A sample of at least 10% of the abstract lists were double-sifted by a senior research fellow and discrepancies rectified, with committee input

	duplicate screening / selection / analysis	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 (OVID), Embase (OVID and the Cochrane Library (Wiley)
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 Appendix D: 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.
XXI	Confidence in	For details please see sections 6.4 and 9.1 of Developing NICE

V	cumulative evidence	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 15: Health economic review protocol

Review question	All questions – health economic evidence
Objectives	To identify health economic studies relevant to any of the review questions.
Search criteria	• Populations, interventions and comparators must be as specified in the clinical review protocol above. • Studies must be of a relevant health economic study design (cost–utility analysis, cost-effectiveness analysis, cost–benefit analysis, cost–consequences analysis, comparative cost analysis). • Studies must not be a letter, editorial or commentary, or a review of health economic evaluations. (Recent reviews will be ordered although not reviewed. The bibliographies will be checked for relevant studies, which will then be ordered.) • Unpublished reports will not be considered unless submitted as part of a call for evidence. • Studies must be in English.
Search strategy	A health economic study search will be undertaken using population-specific terms and a health economic study filter – see appendix B below.
Review strategy	Studies not meeting any of the search criteria above will be excluded. Studies published before 2003, abstract-only studies and studies from non-OECD countries or the USA will also be excluded. Each remaining study will be assessed for applicability and methodological limitations using the NICE economic evaluation checklist which can be found in appendix H of Developing NICE guidelines: the manual (2014).[97] **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 16: 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 17: 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 vs surgery vs radioactive iodine)

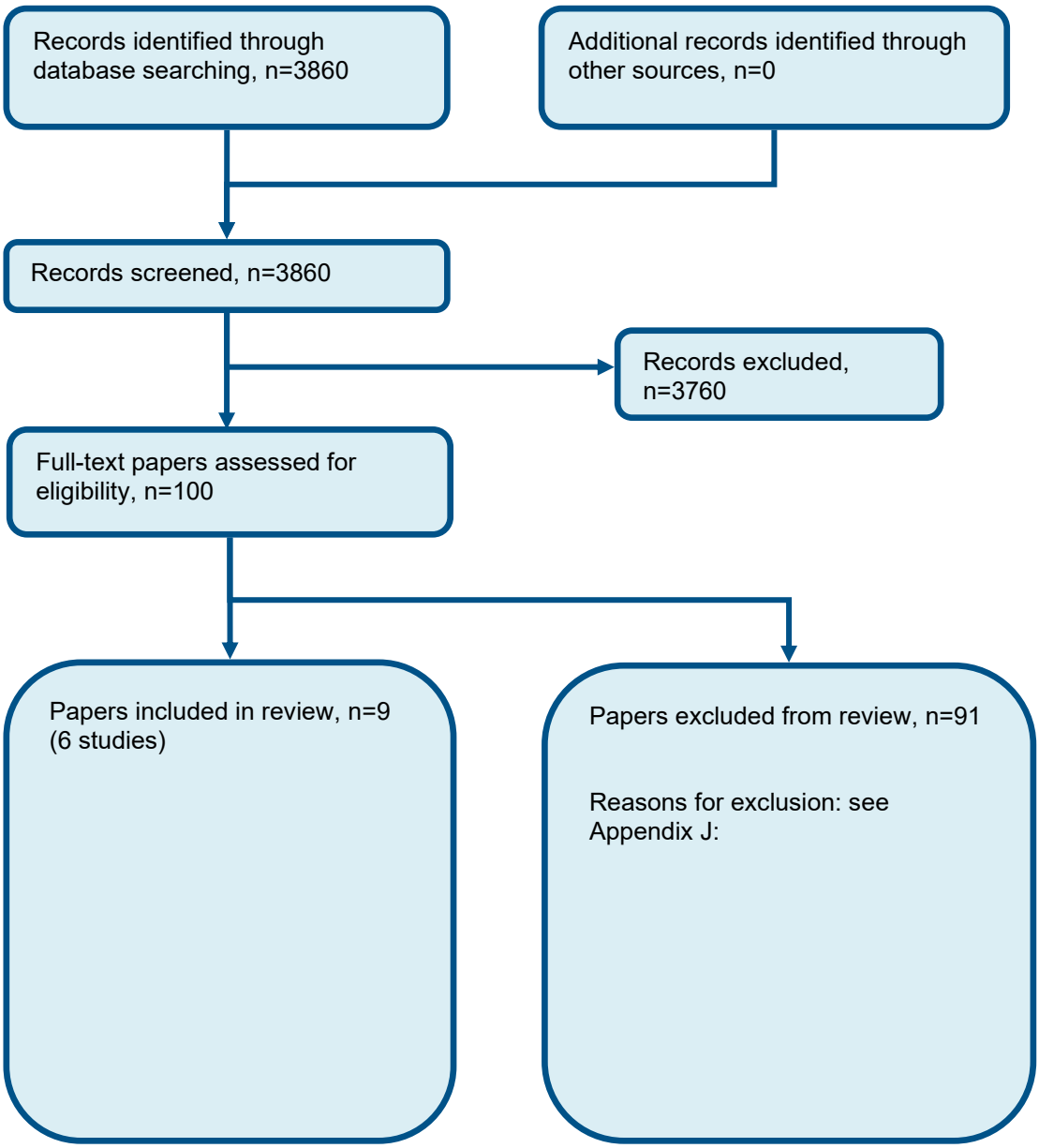
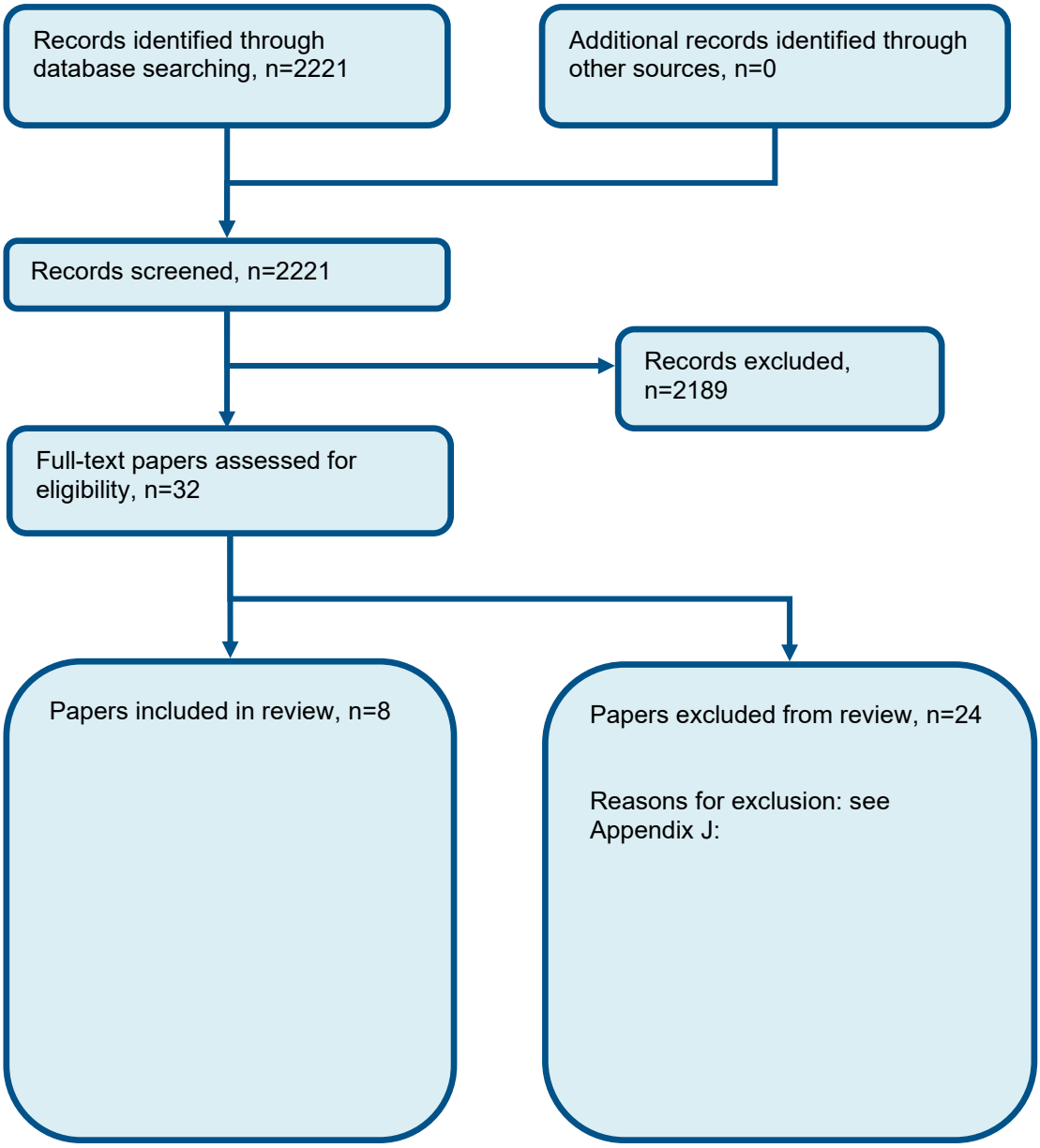


Figure 2: Flow chart of clinical study selection for the review of radioactive iodine safety



Appendix D: Clinical evidence tables

D.1 Drugs vs Surgery vs Radioactive Iodine

Study	Azizi 2005[11]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=104)
Countries and setting	Conducted in Iran; Setting: Not specified
Line of therapy	2nd line
Duration of study	Intervention + follow up: 10 years
Method of assessment of guideline condition	Adequate method of assessment/diagnosis:
Stratum	Failed first line treatment
Subgroup analysis within study	Not applicable
Inclusion criteria	Older than 40, diffuse toxic goitre (Graves'), treated to euthyroidism with MMI for at least 18 months, relapse to hyperthyroidism within 1 year of discontinuation
Exclusion criteria	Did not accept randomisation
Recruitment/selection of patients	Not specified
Age, gender and ethnicity	Age - Mean (SD): 48 (6). Gender (M:F): 69/16. Ethnicity: Not stated
Further population details	1. Age: 18-50 2. Gender: Not stated / Unclear
Indirectness of population	No indirectness

Interventions	(n=52) Intervention 1: Antithyroid drugs. MMI, 10mg twice daily for first month, 10mg daily during second month, maintenance of 2.5-10mg daily from third month on, no discontinuation specified. Duration 10 years. Concurrent medication/care: Usual care (n=52) Intervention 2: Radioactive iodine. Calculated activity based on thyroid weight and iodine uptake, mean dose delivered 7.9 mCi. Duration 10 years. Concurrent medication/care: Usual care. Indirectness: No indirectness
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: RADIOACTIVE IODINE versus ANTITHYROID DRUGS Protocol outcome 1: Euthyroidism - Actual outcome for Failed first line treatment: Euthyroidism at end of follow-up at 10 years; Group 1: 16/41, Group 2: 26/34 Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 11, Reason: 10 lost to follow-up, 1 did not accept randomisation; Group 2 Number missing: 24, Reason: 6 lost to follow-up, 18 did not accept randomisation Protocol outcome 2: Hypothyroidism - Actual outcome for Failed first line treatment: Hypothyroidism at end of follow-up at 10 years; Group 1: 25/41, Group 2: 1/28 Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 11, Reason: 10 lost to follow-up, 1 did not accept randomisation; Group 2 Number missing: 24, Reason: 6 lost to follow-up, 18 did not accept randomisation Protocol outcome 3: Relapse of hyperthyroidism - Actual outcome for Failed first line treatment: Hyperthyroidism at end of follow-up at 10 years; Group 1: 0/41, Group 2: 1/28 Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 11, Reason: 10 lost to follow-up, 1 did not accept randomisation; Group 2 Number missing: 24, Reason: 6 lost to follow-up, 18 did not accept randomisation Protocol outcome 4: Agranulocytosis - Actual outcome for Failed first line treatment: Agranulocytosis at 10 years; Group 1: 0/41, Group 2: 0/28 Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 11, Reason: 10 lost to follow-up, 1 did not accept randomisation; Group 2 Number missing: 24, Reason: 6 lost to follow-up, 18 did not accept randomisation	
Protocol outcomes not reported by the study	Quality of life ; Mortality ; Thyroid ophthalmopathy ; Ischaemic heart disease ; Heart failure ; Arrhythmia ; Osteoporosis ; Impaired cognitive function ; Growth ; Pain ; Symptom scores ; Experience of care ; Healthcare

contacts ; Recurrent laryngeal nerve damage ; Hypocalcaemia ; Hypoparathyroidism ; Bleeding ; Infection ; Liver failure ; Minor drug related adverse events ; Teratogenesis ; Infertility ; Malignancy ; Thyrotoxic storm

Study	Bartalena 1998[16]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=443)
Countries and setting	Conducted in Italy; Setting: Not specified
Line of therapy	1st line
Duration of study	Intervention + follow up: 2.5 years
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Treatment naive/general population
Subgroup analysis within study	Not applicable
Inclusion criteria	Graves' disease, mild or no ophthalmopathy
Exclusion criteria	Severe ophthalmopathy, large goitres, CI to glucocorticoid treatment
Recruitment/selection of patients	Not specified
Age, gender and ethnicity	Age - Mean (range): 42 (15-85). Gender (M:F): 20:80. Ethnicity:
Further population details	1. Age: 18-50 2. Gender: Not applicable
Extra comments	~50% with ophthalmopathy
Indirectness of population	No indirectness
Interventions	(n=150) Intervention 1: Radioactive iodine. MMI was discontinued 5 days before administration of RAI, with dose of 120-150uCi per gram of thyroid tissue, if hypo or hyperthyroid after treatment - corrected with levothyroxine/MMI as appropriate. Duration 1 year. Concurrent medication/care: All given MMI for 3 to 4 months (70% had been given prior to trial achieving euthyroidism in roughly 1/3rd of the 70%). Indirectness: No indirectness (n=148) Intervention 2: Antithyroid drugs. Methimazole given at lowest dose that achieved euthyroidism, no discontinuation specified. Duration 1 year. Concurrent medication/care: Usual care. Indirectness: No indirectness
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: RADIOACTIVE IODINE versus ANTITHYROID DRUGS

Protocol outcome 1: Thyroid ophthalmopathy

- Actual outcome for Treatment naive/general population: Development or worsening of thyroid ophthalmopathy at 1 year; Group 1: 23/150, Group 2: 4/148 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: ; Group 2 Number missing: Protocol outcome 2: Euthyroidism - Actual outcome for Treatment naive/general population: Euthyroidism at end of follow-up (including RAI patients requiring levothyroxine/MMI) at 1 year; Group 1: 128/150, Group 2: 145/148 Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: ; Group 2 Number missing: Protocol outcome 3: Hypothyroidism - Actual outcome for Treatment naive/general population: Hypothyroidism at end of follow-up (including RAI patients requiring levothyroxine/MMI) at 1 year; Group 1: 20/150, Group 2: 2/148 Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: ; Group 2 Number missing: - Actual outcome for Treatment naive/general population: Hyperthyroidism at end of follow-up (including RAI patients requiring levothyroxine/MMI) at 1 year; Group 1: 2/150, Group 2: 1/148 Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: ; Group 2 Number missing:	
Protocol outcomes not reported by the study	Quality of life ; Mortality ; Relapse of hyperthyroidism ; Ischaemic heart disease ; Heart failure ; Arrhythmia ; Osteoporosis ; Impaired cognitive function ; Growth ; Pain ; Symptom scores ; Experience of care ; Healthcare contacts ; Recurrent laryngeal nerve damage ; Hypocalcaemia ; Hypoparathyroidism ; Bleeding ; Infection ; Agranulocytosis ; Liver failure ; Minor drug related adverse events ; Teratogenesis ; Infertility ; Malignancy ; Thyrotoxic storm

Study	Chen 2009[28]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=460)
Countries and setting	Conducted in China; Setting: Not stated
Line of therapy	1st line
Duration of study	Intervention + follow up: 9 years
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Newly diagnosed hyperthyroidism, no previous thyroid treatment, 24 hour uptake of 131I >40%
Exclusion criteria	Severe liver or kidney damage, agranulocytosis, pregnancy or lactation, less than 8 years of age
Recruitment/selection of patients	Screened 2021, excluding 1519 with previous treatment, others refused or exclusion criteria
Age, gender and ethnicity	Age - Mean (SD): 37 (14). Gender (M:F): 33:67. Ethnicity:
Further population details	1. Age: 18-50 2. Gender: Systematic review: mixed
Extra comments	75% GD, 23% MNTG, 2% UNTG
Indirectness of population	No indirectness
Interventions	(n=230) Intervention 1: Antithyroid drugs. Either MMI or PTU, for at least 18 months, initial dose based on severity of symptoms and titrated throughout to TSH. If recurrence after withdrawal, reinstated. . Duration 9 years . Concurrent medication/care: All also received propranolol as necessary. Advised to restrict iodine rich foods in diet. Examined 2-4 weekly in first year, 3-6 months thereafter if stable. Indirectness: No indirectness (n=230) Intervention 2: Radioactive iodine. No pre-treatment with ATD. Therapeutic activity from 1.85-4.44MBq per gram thyroid/lesion weight, calculated activity (based on weight, and 24hr iodine uptake). Maximum activity limited to 555MBq. At 3 months and 6 months determined if 2nd (10%) or 3rd (2.5%) treatment required. Duration 9 years. Concurrent medication/care: All also received propranolol as necessary. Advised to restrict iodine rich foods in diet. Examined 2-4 weekly in first year, 3-6 months thereafter if stable. Indirectness: No indirectness
Funding	Academic or government funding

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: RADIOACTIVE IODINE versus MMI/PTU

Protocol outcome 1: Mortality
- Actual outcome for Treatment naive/general population: Mortality at 9 years; Group 1: 0/209, Group 2: 0/177
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low;
Indirectness of outcome: No indirectness ; Group 1 Number missing: 21, Reason: Loss to follow up; Group 2 Number missing: 53, Reason: 16 loss to follow-up, 37 excluded due to AEs

Protocol outcome 2: Thyroid ophthalmopathy
- Actual outcome for Treatment naive/general population: New cases of thyroid ophthalmopathy at 9 years; Group 1: 26/151, Group 2: 14/138
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low;
Indirectness of outcome: No indirectness ; Group 1 Number missing: 21, Reason: Loss to follow up; Group 2 Number missing: 53, Reason: 16 loss to follow-up, 37 excluded due to AEs

Protocol outcome 3: Euthyroidism
- Actual outcome for Treatment naive/general population: Normal T3+T4, no medication required at 9 years; Group 1: 146/209, Group 2: 73/177
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low;
Indirectness of outcome: No indirectness ; Group 1 Number missing: 21, Reason: Loss to follow up; Group 2 Number missing: 53, Reason: 16 loss to follow-up, 37 excluded due to AEs

Protocol outcome 4: Hypothyroidism
- Actual outcome for Treatment naive/general population: Clinical hypothyroidism, abnormal T3/T4 and TSH at 9 years; Group 1: 19/209, Group 2: 6/177
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low;
Indirectness of outcome: No indirectness ; Group 1 Number missing: 21, Reason: Loss to follow up; Group 2 Number missing: 53, Reason: 16 loss to follow-up, 37 excluded due to AEs

Protocol outcome 5: Relapse of hyperthyroidism
- Actual outcome for Treatment naive/general population: Relapse or persistence of hyperthyroidism, abnormal T3/T4 or TSH at 9 years; Group 1: 18/209, Group 2: 88/177
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low;
Indirectness of outcome: No indirectness ; Group 1 Number missing: 21, Reason: Loss to follow up; Group 2 Number missing: 53, Reason: 16 loss to follow-up, 37 excluded due to AEs

Protocol outcome 6: Agranulocytosis
- Actual outcome for Treatment naive/general population: Agranulocytosis at 9 years; Group 1: 0/209, Group 2: 7/214
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low;

Indirectness of outcome: No indirectness ; Group 1 Number missing: 21, Reason: Loss to follow up; Group 2 Number missing: 16, Reason: Loss to follow up Protocol outcome 7: Liver failure - Actual outcome for Treatment naive/general population: Severe liver damage at 9 years; Group 1: 0/209, Group 2: 5/214 Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 21, Reason: Loss to follow up; Group 2 Number missing: 16, Reason: Loss to follow up Protocol outcome 8: Malignancy - Actual outcome for Treatment naive/general population: Malignancy at 9 years; Group 1: 0/209, Group 2: 0/177 Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 21, Reason: Loss to follow up; Group 2 Number missing: 53, Reason: 16 loss to follow-up, 37 excluded due to AEs - Actual outcome for Treatment naive/general population: Thyroid storm at 9 years; Group 1: 0/209, Group 2: 0/177 Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 21, Reason: Loss to follow up; Group 2 Number missing: 53, Reason: 16 loss to follow-up, 37 excluded due to AEs	
Protocol outcomes not reported by the study	Quality of life ; Ischaemic heart disease ; Heart failure ; Arrhythmia ; Osteoporosis ; Impaired cognitive function ; Growth ; Pain ; Symptom scores ; Experience of care ; Healthcare contacts ; Recurrent laryngeal nerve damage ; Hypocalcaemia ; Hypoparathyroidism ; Bleeding ; Infection ; Minor drug related adverse events ; Teratogenesis ; Infertility ; Thyrotoxic storm

Study	Kansara 2017[70]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=60)
Countries and setting	Conducted in India; Setting: Tertiary level referral centre
Line of therapy	1st line
Duration of study	Intervention + follow up: 1 year
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	20-50 years old, treatment naive
Exclusion criteria	History of thyroid disease, significant ophthalmopathy (clinical activity score >1), malignancy, previous exposure to RAI, known systemic disorders, long term use of corticosteroids or insulin
Age, gender and ethnicity	Age - Mean (SD): 33 (4.2). Gender (M:F): Not stated. Ethnicity:
Further population details	1. Age: 18-50 2. Gender: Not stated / Unclear
Indirectness of population	No indirectness
Interventions	(n=30) Intervention 1: Antithyroid drugs. Carbimazole, 30mg initially for 2 months, tapering as per clinical status. Duration 1 year. Concurrent medication/care: Usual care. Indirectness: No indirectness (n=30) Intervention 2: Radioactive iodine. Orally, single dose of 131I 10mCi, capsule form with water, . Duration 1 year. Concurrent medication/care: Usual care. Indirectness: No indirectness
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: RADIOACTIVE IODINE versus CARBIMAZOLE

Protocol outcome 1: Euthyroidism
- Actual outcome for Treatment naive/general population: Biochemical and clinical euthyroidism at 1 year; Group 1: 4/28, Group 2: 22/29
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low;
Indirectness of outcome: No indirectness ; Group 1 Number missing: 2, Reason: Lost to follow-up; Group 2 Number missing: 1, Reason: Lost to follow-up
- Actual outcome for Treatment naive/general population: Clinical hypothyroidism at 1 year; Group 1: 24/28, Group 2: 2/29

Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2, Reason: Lost to follow-up; Group 2 Number missing: 1, Reason: Lost to follow-up Protocol outcome 2: Relapse of hyperthyroidism - Actual outcome for Treatment naive/general population: Relapse/persistent hyperthyroidism at 1 year; Group 1: 0/28, Group 2: 0/29 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2, Reason: Lost to follow-up; Group 2 Number missing: 1, Reason: Lost to follow-up	
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 ; Recurrent laryngeal nerve damage ; Hypocalcaemia ; Hypoparathyroidism ; Bleeding ; Infection ; Agranulocytosis ; Liver failure ; Minor drug related adverse events ; Teratogenesis ; Infertility ; Malignancy ; Thyrotoxic storm

Study (subsidiary papers)	Törring 1996[127] (Abraham-nordling 2005[1], Tallstedt 1992[124])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=179)
Countries and setting	Conducted in Sweden; Setting: Not specified
Line of therapy	1st line
Duration of study	Intervention + follow up: Maximum 21 years
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Treatment naive/general population
Subgroup analysis within study	Not applicable
Inclusion criteria	Graves' disease
Exclusion criteria	Previous thyroid disease
Recruitment/selection of patients	Not specified
Age, gender and ethnicity	Age - Mean (SD): Younger group = 29 (4), older group 45 (6). Gender (M:F): 16:84. Ethnicity:
Further population details	1. Age: Not applicable 2. Gender: Not applicable
Indirectness of population	No indirectness
Interventions	(n=71) Intervention 1: Antithyroid drugs. 10mg MMI 4x daily for 18 months, thyroxine 0.1 to 0.3mg daily after 3-5 weeks to provide normal T3 and low TSH. Beta blockers given for initial weeks. Examined monthly for 2 months after initiation, then 3 monthly. After discontinuation examined twice in first year, once yearly. Duration Max 21 years follow-up. Concurrent medication/care: Usual care. Indirectness: No indirectness (n=67) Intervention 2: Surgery. Beta blockers before surgery for ~1 month, bilateral subtotal thyroidectomy, leaving posterior capsule and 1g or less of each lobe, thyroxine 0.1 to 0.3mg daily afterwards, seen after 5 weeks and then every 3 months during 1st year after surgery and once yearly thereafter. Duration Max 21 years follow-up. Concurrent medication/care: Usual care. Indirectness: No indirectness (n=41) Intervention 3: Radioactive iodine. First single oral dose of iodine 131I, dose based on size of thyroid, uptake and half-life aiming at 120Gy dose delivered. Beta blockers also given unless CI. 18 patients needed more than 1 one dose of RAI. Duration Max 21 years follow-up. Concurrent medication/care: Usual care. Indirectness: No indirectness
Funding	Academic or government funding

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SURGERY versus ANTITHYROID DRUGS

Protocol outcome 1: Thyroid ophthalmopathy
- Actual outcome for Treatment naive/general population: New or worsening ophthalmopathy at ~4 years follow-up; Group 1: 9/64, Group 2: 8/65
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low;
Indirectness of outcome: No indirectness ; Group 1 Number missing: -, Reason: Not fully specified across publications; Group 2 Number missing: -, Reason: Not fully specified across publications

Protocol outcome 2: Relapse of hyperthyroidism
- Actual outcome for Treatment naive/general population: Relapse or persistence of hyperthyroidism at ~4 years follow-up; Group 1: 4/65, Group 2: 26/68
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low;
Indirectness of outcome: No indirectness ; Group 1 Number missing: 2, Reason: Did not have surgery; Group 2 Number missing: 3, Reason: Did not comply/randomisation error

Protocol outcome 3: Osteoporosis
- Actual outcome for Treatment naive/general population: Osteoporosis (self-reported) at 14-21 years follow-up; Group 1: 8/56, Group 2: 5/55
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low;
Indirectness of outcome: No indirectness ; Group 1 Number missing: -, Reason: Not fully specified across publications; Group 2 Number missing: -, Reason: Not fully specified across publications

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: RADIOACTIVE IODINE versus ANTITHYROID DRUGS

Protocol outcome 1: Thyroid ophthalmopathy
- Actual outcome for Treatment naive/general population: New or worsening ophthalmopathy at ~4 years follow-up; Group 1: 13/39, Group 2: 4/38
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low;
Indirectness of outcome: No indirectness ; Group 1 Number missing: -, Reason: Not fully specified across publications; Group 2 Number missing: -, Reason: Not fully specified across publications

Protocol outcome 2: Relapse of hyperthyroidism
- Actual outcome for Treatment naive/general population: Relapse or persistence of hyperthyroidism at ~4 years follow-up; Group 1: 8/39, Group 2: 16/38
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low;
Indirectness of outcome: No indirectness ; Group 1 Number missing: 2, Reason: Rejected assignment; Group 2 Number missing: 3, Reason: Did not comply/randomisation error

Protocol outcome 3: Osteoporosis

- Actual outcome for Treatment naive/general population: Osteoporosis (self-reported) at 14-21 years follow-up; Group 1: 6/34, Group 2: 5/36 Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: -, Reason: Not fully specified across publications; Group 2 Number missing: -, Reason: Not fully specified across publications RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: RADIOACTIVE IODINE versus SURGERY Protocol outcome 1: Thyroid ophthalmopathy - Actual outcome for Treatment naive/general population: New or worsening ophthalmopathy at ~4 years follow-up; Group 1: 13/39, Group 2: 6/37 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: -, Reason: Not fully specified across publications; Group 2 Number missing: -, Reason: Not fully specified across publications Protocol outcome 2: Relapse of hyperthyroidism - Actual outcome for Treatment naive/general population: Relapse or persistence of hyperthyroidism at ~4 years follow-up; Group 1: 8/39, Group 2: 3/37 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2, Reason: Rejected assignment; Group 2 Number missing: 0 Protocol outcome 3: Osteoporosis - Actual outcome for Treatment naive/general population: Osteoporosis (self-reported) at 14-21 years follow-up; Group 1: 6/34, Group 2: 7/34 Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: -, Reason: Not fully specified across publications; Group 2 Number missing: -, Reason: Not fully specified across publications	
Protocol outcomes not reported by the study	Quality of life ; Mortality ; Euthyroidism ; Hypothyroidism ; Ischaemic heart disease ; Heart failure ; Arrhythmia ; Impaired cognitive function ; Growth ; Pain ; Symptom scores ; Experience of care ; Healthcare contacts ; Recurrent laryngeal nerve damage ; Hypocalcaemia ; Hypoparathyroidism ; Bleeding ; Infection ; Agranulocytosis ; Liver failure ; Minor drug related adverse events ; Teratogenesis ; Infertility ; Malignancy ; Thyrotoxic storm

Study (subsidiary papers)	Träisk 2009[128] (Abraham-nordling 2010[3])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=313)
Countries and setting	Conducted in Sweden; Setting: Sweden, outpatients for RAI
Line of therapy	1st line
Duration of study	Intervention + follow up: 4 years
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Treatment naive/general population
Subgroup analysis within study	Not applicable
Inclusion criteria	35-69 years old, symptomatic Graves' disease, activity of oral dose of radioactive iodine </= 600MBq
Exclusion criteria	Previous treatment with thyroid drugs/surgery/radioactive iodine , severe ophthalmopathy, incipient toxic crisis, coronary heart disease, pregnancy, breast-feeding, pregnancy planned within 2 years
Recruitment/selection of patients	Not specified
Age, gender and ethnicity	Age - Mean (SD): 51 (8). Gender (M:F): 11:89. Ethnicity: Not stated
Further population details	1. Age: 50-65 2. Gender: Not applicable
Extra comments	Ophthalmopathy at baseline in 13%
Indirectness of population	No indirectness
Interventions	(n=163) Intervention 1: Radioactive iodine. Beta blocker pre-treatment, aim for one dose, calculated activity based on mass, estimated uptake and effective half-life. Duration 4 years. Concurrent medication/care: Usual care. Indirectness: No indirectness (n=150) Intervention 2: Antithyroid drugs. MMI given 15mg twice daily for 2 weeks, then 50ug of thyroxine added and increased to 100ug 2 weeks later. At 6 weeks adjusted to normalise T3/T4 and bring TSH to less than 0.4mIU/litre. Beta blockers used for symptomatic treatment. MMI replaced by PTU in people showing serious adverse reactions. Discontinued after 18 months, levothyroxine continued for 1 more month. Duration 4 years. Concurrent medication/care: Usual care. Indirectness: No indirectness
Funding	Academic or government funding

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: RADIOACTIVE IODINE versus ANTITHYROID DRUGS Protocol outcome 1: Thyroid ophthalmopathy - Actual outcome for Treatment naive/general population: Relapse of hyperthyroidism at 3 years; Group 1: 2/147, Group 2: 33/137 Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 16, Reason: Lost to follow-up; Group 2 Number missing: 10, Reason: Lost to follow-up	
Protocol outcomes not reported by the study	Quality of life ; Mortality ; Euthyroidism ; Hypothyroidism ; Relapse of hyperthyroidism ; Ischaemic heart disease ; Heart failure ; Arrhythmia ; Osteoporosis ; Impaired cognitive function ; Growth ; Pain ; Symptom scores ; Experience of care ; Healthcare contacts ; Recurrent laryngeal nerve damage ; Hypocalcaemia ; Hypoparathyroidism ; Bleeding ; Infection ; Agranulocytosis ; Liver failure ; Minor drug related adverse events ; Teratogenesis ; Infertility ; Malignancy ; Thyrotoxic storm

D.2 Radioactive Iodine safety

Study	**Franklyn 1999[38]**
Study type	Non randomised study

Number of studies (number of participants)	1 (n=7417)
Countries and setting	Conducted in United Kingdom; Setting: Nil else stated
Line of therapy	1st line
Duration of study	--:
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Fixed dose <400MBq
Subgroup analysis within study	Not applicable
Inclusion criteria	Treated with RAI in WM in UK between 1950 and 1991, did not die before 1971, found on register with ONS, not emigrated, registered with GP
Exclusion criteria	Nil else stated
Recruitment/selection of patients	Nil else stated
Age, gender and ethnicity	Age - Mean (SD): 56.6 (12.7). Gender (M:F): 17:83. Ethnicity: Not stated
Further population details	1. Age: 2. Gender:
Extra comments	Nil else stated
Indirectness of population	No indirectness
Interventions	(n=7417) Intervention 1: RAI - RAI alone. Mean dose 308Mbq (SD 232). Duration Mean follow-up 9.7 years. Concurrent medication/care: Not specified . Indirectness: No indirectness (n=7417) Intervention 2: General population. Age, sex and period matched SIR from UK regional cancer registries. Duration Mean follow-up 9.7 years. Concurrent medication/care: Nil else stated. Indirectness: No indirectness
Funding	Other (Government + BUPA foundation)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: RAI ALONE versus GENERAL POPULATION

Protocol outcome 1: Total cancer diagnoses
- Actual outcome for Fixed dose <400MBq: All cancer diagnoses at Mean follow-up 9.7 years; RR; 0.83 (95%CI 0.77 to 0.9);
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low;
Indirectness of outcome: No indirectness ; Group 1 Number missing: ; Group 2 Number missing:

Protocol outcomes not reported by the study	Quality of life ; Cancer diagnoses in iodine uptake glands ; Cancer diagnoses in non-iodine uptake glands ; Infertility

Study	Franklyn 2005[40]
Study type	Non randomised study
Number of studies (number of participants)	1 (n=2668)
Countries and setting	Conducted in United Kingdom; Setting: Nil else
Line of therapy	1st line
Duration of study	Intervention + follow up: Median 5.6 years
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Fixed dose <400MBq
Subgroup analysis within study	Not applicable
Inclusion criteria	>40, hyperthyroidism, treated in West Midlands with radioiodine between 1984 and 2002, records available
Exclusion criteria	Nil else
Recruitment/selection of patients	Not specified
Age, gender and ethnicity	Age - Median (range): 62 (40 to >80). Gender (M:F): 19:81. Ethnicity: Not stated
Further population details	1. Age: 2. Gender:
Indirectness of population	No indirectness
Interventions	(n=2668) Intervention 1: RAI - RAI alone. Fixed dose, either 185 or 370MBq, 84.3% received one dose only. Duration Median follow-up 5.6 years. Concurrent medication/care: Nil else stated. Indirectness: No indirectness (n=2668) Intervention 2: General population. From WHO databank, age, sex and year matched cohort. Duration Median follow-up 5.6 years. Concurrent medication/care: Nil else stated. Indirectness: No indirectness
Funding	Study funded by industry (Some funding from BUPA)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: RAI ALONE versus GENERAL POPULATION

Protocol outcome 1: Total cancer diagnoses
- Actual outcome for Fixed dose <400MBq: Cancer mortality at 5.6 years; RR; 0.99 (95%CI 0.82 to 1.2);
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low;
Indirectness of outcome: Serious indirectness ; Key confounders: Age, sex and year matched SMR; Group 1 Number missing: ; Group 2 Number missing:

Quality of life ; Cancer diagnoses in iodine uptake glands ; Cancer diagnoses in non-iodine uptake glands ; Infertility	
Protocol outcomes not reported by the study	

Study	Giesecke 2018[42]
Study type	Non randomised study
Number of studies (number of participants)	1 (n=10992)
Countries and setting	Conducted in Sweden; Setting: Nil else
Line of therapy	1st line
Duration of study	Intervention + follow up: Mean 16.3 years
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Cause not specified
Subgroup analysis within study	Not applicable
Inclusion criteria	RAI at Karolinska University Hospital or surgery in Stockholm between 1976 and 2013, older than 35, certain aetiology of hyperthyroidism, not treated with both RAI and surgery
Exclusion criteria	Nil else
Recruitment/selection of patients	Nil else
Age, gender and ethnicity	Age - Other: Mean for RAI 64, mean for surgery 47. Gender (M:F): 15:85. Ethnicity: Not stated
Further population details	1. Age: 2. Gender:
Extra comments	50% Graves disease in RAI arm, 63% in surgery arm
Indirectness of population	No indirectness
Interventions	(n=10250) Intervention 1: RAI - RAI alone. Dose not stated. Duration 16.3 years. Concurrent medication/care: Nil else stated. Indirectness: No indirectness (n=742) Intervention 2: ATD/SUR - SUR. No details provided. Duration 16.3 years. Concurrent medication/care: No details provided . Indirectness: No indirectness
Funding	Academic or government funding
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: RAI ALONE versus SUR Protocol outcome 1: Total cancer diagnoses - Actual outcome for Fixed dose <400MBq: Cancer mortality at 16.3 years follow-up: HR: 0.96 (95%CI 0.73 to 1.26);	

Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: ; Group 2 Number missing:	
Protocol outcomes not reported by the study	Quality of life ; Cancer diagnoses in iodine uptake glands ; Cancer diagnoses in non-iodine uptake glands ; Infertility

Study	Goldman 1988[44]
Study type	Non randomised study
Number of studies (number of participants)	1 (n=1762)
Countries and setting	Conducted in USA; Setting: None stated
Line of therapy	1st line
Duration of study	Intervention + follow up: 17.2 years
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Cause not specified
Subgroup analysis within study	Not applicable
Inclusion criteria	Women with hyperthyroidism treated at MGH between 1946 and 1964 with I131
Exclusion criteria	None stated
Recruitment/selection of patients	None stated
Age, gender and ethnicity	Age - Other: Not stated. Gender (M:F): All women. Ethnicity: Not stated
Further population details	1. Age: 2. Gender:
Indirectness of population	No indirectness
Interventions	(n=1762) Intervention 1: RAI - RAI alone. None stated. Duration 17.2 years. Concurrent medication/care: None stated. Indirectness: No indirectness (n=1762) Intervention 2: General population. Age, sex, race, year matched incidence from state cancer register. Duration 17.2 years. Concurrent medication/care: None stated. Indirectness: No indirectness
Funding	Academic or government funding

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: RAI ALONE versus GENERAL POPULATION

Protocol outcome 1: Total cancer diagnoses
- Actual outcome for Cause not specified: Total cancer diagnoses, SIR at 17.2 years; RR; 0.8 (95%CI 0.6 to 1.1);
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low;
Indirectness of outcome: No indirectness ; Group 1 Number missing: ; Group 2 Number missing:

Quality of life ; Cancer diagnoses in iodine uptake glands ; Cancer diagnoses in non-iodine uptake glands ; Infertility	
Protocol outcomes not reported by the study	

Study	Hoffman 1982[59]
Study type	Non randomised study
Number of studies (number of participants)	1 (n=3146)
Countries and setting	Conducted in USA; Setting: Nil else stated
Line of therapy	1st line
Duration of study	Intervention + follow up: 15 years mean for RAI, 21 years mean for surgery
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Fixed dose <400MBq
Subgroup analysis within study	Not applicable
Inclusion criteria	White, female, treated for hyperthyroidism at Mayo clinic between 1946 and 1964, confirmed diagnosis of hyperthyroidism, no other isotope treatment, resident of USA
Exclusion criteria	Nil else stated
Recruitment/selection of patients	Nil else stated
Age, gender and ethnicity	Age - Other: Mean age at Tx 56.8 for RAI, 45.7 for surgery. Gender (M:F): Only women. Ethnicity: Only white patients
Further population details	1. Age: 2. Gender:
Extra comments	73% mild-moderate disease, ~50% Graves disease
Indirectness of population	No indirectness
Interventions	(n=1005) Intervention 1: RAI - RAI alone. Mean number of treatments 1.2, mean dose 10.6mCi (~392 MBq). Duration Mean 15 years follow-up. Concurrent medication/care: Nil else stated. Indirectness: No indirectness (n=2141) Intervention 2: ATD/SUR - SUR. Nil else stated. Duration Mean 21 years follow-up. Concurrent medication/care: Nil else stated. Indirectness: No indirectness
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: RAI ALONE versus SUR

Protocol outcome 1: Total cancer diagnoses
- Actual outcome for Fixed dose <400MBq: Cancer incidence at all sites, adjusted for age, year of treatment and duration of follow-up at Mean follow-up 15 years for RAI.

21 years for surgery; RR; 1.0 (95%CI 0.7 to 1.3); Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - High, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: ; Group 2 Number missing:	
Protocol outcomes not reported by the study	Quality of life ; Cancer diagnoses in iodine uptake glands ; Cancer diagnoses in non-iodine uptake glands ; Infertility

Study	Holm 1991[62]
Study type	Non randomised study
Number of studies (number of participants)	1 (n=10207)
Countries and setting	Conducted in Sweden; Setting: Not stated
Line of therapy	1st line
Duration of study	Intervention + follow up: Mean 15 years
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Fixed dose >400MBq
Subgroup analysis within study	Not applicable
Inclusion criteria	Under 75, treated for hyperthyroidism with RAI at one of 7 departments in Sweden, sufficient information on names and DoB
Exclusion criteria	Not stated
Recruitment/selection of patients	Not stated
Age, gender and ethnicity	Age - Mean (range): 57 (13-74). Gender (M:F): Not stated. Ethnicity: Not stated
Further population details	1. Age: 2. Gender:
Extra comments	51% with Graves' disease
Indirectness of population	No indirectness
Interventions	(n=10207) Intervention 1: RAI - RAI alone. Mean dose 506MBq, 59% received one treatment. Duration 15 years. Concurrent medication/care: Not stated. Indirectness: No indirectness (n=10207) Intervention 2: General population. Age, sex, region and year matched SIR based on Swedish Cancer register. Duration 15 years. Concurrent medication/care: Not stated. Indirectness: No indirectness
Funding	Academic or government funding

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: RAI ALONE versus GENERAL POPULATION

Protocol outcome 1: Total cancer diagnoses
- Actual outcome for Fixed dose >400MBq: Total cancer diagnoses. SIR at 15 year follow-up: RR: 1.10 (95%CI 1.02 to 1.17);

Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: ; Group 2 Number missing:	
Protocol outcomes not reported by the study	Quality of life ; Cancer diagnoses in iodine uptake glands ; Cancer diagnoses in non-iodine uptake glands ; Infertility

Study	Metso 2007[90]
Study type	Non randomised study
Number of studies (number of participants)	1 (n=2793)
Countries and setting	Conducted in Finland; Setting: Nil else stated
Line of therapy	1st line
Duration of study	Intervention + follow up: 10 years
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Fixed dose <400MBq
Subgroup analysis within study	Not applicable
Inclusion criteria	Treated with RAI for hyperthyroidism at Tampere hospital between 1965 and 2002
Exclusion criteria	Nil else stated
Recruitment/selection of patients	Nil else stated
Age, gender and ethnicity	Age - Median (range): 62 (50-75). Gender (M:F): 16:84. Ethnicity: Not stated
Further population details	1. Age: 2. Gender:
Indirectness of population	No indirectness
Interventions	(n=2793) Intervention 1: RAI - RAI alone. Mean dose 305MBq, 80.3% received a single dose. Duration 9.8 years follow-up. Concurrent medication/care: Nil else stated (n=2793) Intervention 2: General population. Age and sex matched control selected from Population register. Duration 9.8 years follow-up. Concurrent medication/care: Nil else stated
Funding	Academic or government funding

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: RAI ALONE versus GENERAL POPULATION

Protocol outcome 1: Total cancer diagnoses
- Actual outcome for Fixed dose <400MBq: Cancer, all diagnoses SIR at 10 years follow-up; RR; 1.25 (95%CI 1.08 to 1.46);
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low;
Indirectness of outcome: No indirectness ; Group 1 Number missing: ; Group 2 Number missing:

Quality of life ; Cancer diagnoses in iodine uptake glands ; Cancer diagnoses in non-iodine uptake glands ; Infertility	
Protocol outcomes not reported by the study	

Study	Ryodi 2015[115]
Study type	Non randomised study
Number of studies (number of participants)	1 (n=6148)
Countries and setting	Conducted in Finland; Setting: Not stated
Line of therapy	1st line
Duration of study	Intervention + follow up: Median follow-up 10 years
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Cause not specified
Subgroup analysis within study	Not applicable
Inclusion criteria	People treated with surgery for hyperthyroidism in Finland between 1986 and 2007, people treated with RAI for hyperthyroidism at Tampere University Hospital, reference population randomly chosen from national population register with 3 age and sex matched control subjects
Exclusion criteria	Not stated
Recruitment/selection of patients	Not stated
Age, gender and ethnicity	Age - Median (range): 46 for thyroidectomy, 59 for RAI. Gender (M:F): 16:84. Ethnicity:
Further population details	1. Age: 2. Gender:
Indirectness of population	No indirectness
Interventions	(n=1814) Intervention 1: RAI - RAI alone. No details provided. Duration Median follow-up 10 years. Concurrent medication/care: No details provided. Indirectness: No indirectness (n=4334) Intervention 2: ATD/SUR - SUR. No details provided. Duration Median follow-up 10 years. Concurrent medication/care: No details provided. Indirectness: No indirectness (n=18432) Intervention 3: General population. No details provided. Duration Median follow-up 10 years. Concurrent medication/care: No details provided. Indirectness: No indirectness
Funding	No funding
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: RAI ALONE versus SUR	

Protocol outcome 1: Total cancer diagnoses - Actual outcome for Cause not specified: Total cancer diagnoses at Please enter a time period.; RR; 1.03 (95%CI 0.86 to 1.23); Risk of bias: All domain - Very high, Selection - Very high, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness, Comments: Obtained from Finnish cancer registry which captures 98% of cancer diagnoses, excluding benign, uncertain or borderline tumours; Key confounders: Adjusted for etiology, age, gender; Group 1 Number missing: ; Group 2 Number missing:	
Protocol outcomes not reported by the study	Quality of life ; Cancer diagnoses in iodine uptake glands ; Cancer diagnoses in non-iodine uptake glands ; Infertility

Appendix E: Forest plots

E.1 Drugs vs Surgery vs Radioactive iodine

E.1.1 Radioactive iodine vs antithyroid drugs, adults with Graves' disease, first line treatment

Figure 3: Mortality

Study or Subgroup	RAI Events	RAI Total	ATD Events	ATD Total	Weight	Peto Odds Ratio Peto, Fixed, 95% CI
Chen 2009	0	209	0	177		Not estimable
Total (95% CI)		**209**		**177**		**Not estimable**
Total events	0		0			

Heterogeneity: Not applicable
Test for overall effect: Not applicable

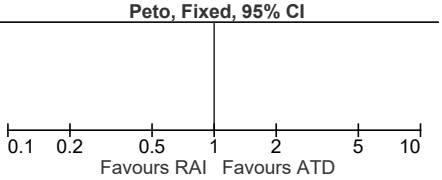


Figure 4: Ophthalmopathy (new or worsening cases)

Study or Subgroup	RAI Events	RAI Total	ATD Events	ATD Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Bartalena 1998	23	150	4	148	7.2%	5.67 [2.01, 16.01]
Chen 2009	26	151	14	138	26.2%	1.70 [0.92, 3.12]
Torring 1996	13	39	4	38	7.3%	3.17 [1.13, 8.85]
Traisk 2009	63	147	32	137	59.3%	1.83 [1.29, 2.62]
Total (95% CI)		**487**		**461**	**100.0%**	**2.17 [1.64, 2.88]**
Total events	125		54			

Heterogeneity: Chi² = 5.30, df = 3 (P = 0.15); I² = 43%
Test for overall effect: Z = 5.37 (P < 0.00001)

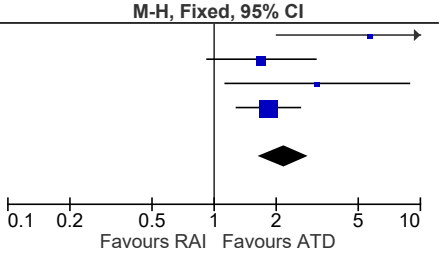


Figure 5: Euthyroidism

Study or Subgroup	RAI Events	RAI Total	ATD Events	ATD Total	Weight	Risk Ratio M-H, Random, 95% CI
Bartalena 1998	128	150	145	148	38.5%	0.87 [0.81, 0.93]
Chen 2009	146	209	73	177	37.6%	1.69 [1.39, 2.06]
Kansara 2017	4	28	22	29	23.8%	0.19 [0.07, 0.48]
Total (95% CI)		**387**		**354**	**100.0%**	**0.78 [0.37, 1.62]**
Total events	278		240			

Heterogeneity: Tau² = 0.36; Chi² = 68.33, df = 2 (P < 0.00001); I² = 97%
Test for overall effect: Z = 0.68 (P = 0.50)

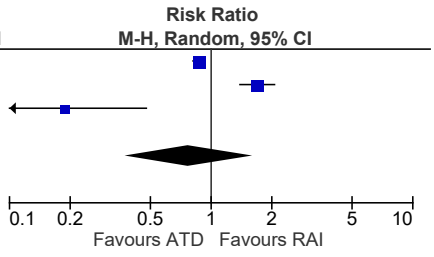


Figure 6: Hypothyroidism

Study or Subgroup	RAI Events	RAI Total	ATD Events	ATD Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Bartalena 1998	20	150	2	148	19.2%	9.87 [2.35, 41.47]
Chen 2009	19	209	6	177	62.0%	2.68 [1.10, 6.57]
Kansara 2017	24	28	2	29	18.8%	12.43 [3.24, 47.74]
Total (95% CI)		**387**		**354**	**100.0%**	**5.89 [3.12, 11.11]**
Total events	63		10			

Heterogeneity: Chi² = 4.64, df = 2 (P = 0.10); I² = 57%
Test for overall effect: Z = 5.48 (P < 0.00001)

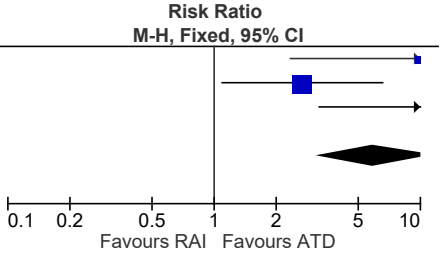


Figure 7: Hyperthyroidism

Study or Subgroup	RAI Events	RAI Total	ATD Events	ATD Total	Weight	Risk Ratio M-H, Random, 95% CI
Bartalena 1998	2	150	1	148	12.1%	1.97 [0.18, 21.53]
Chen 2009	18	209	88	177	34.7%	0.17 [0.11, 0.28]
Kansara 2017	0	28	0	29		Not estimable
Torring 1996	8	39	16	38	31.5%	0.49 [0.24, 1.00]
Traisk 2009	2	147	33	137	21.7%	0.06 [0.01, 0.23]
Total (95% CI)		**573**		**529**	**100.0%**	**0.25 [0.09, 0.69]**
Total events	30		138			

Heterogeneity: Tau² = 0.71; Chi² = 13.09, df = 3 (P = 0.004); I² = 77%
Test for overall effect: Z = 2.67 (P = 0.008)

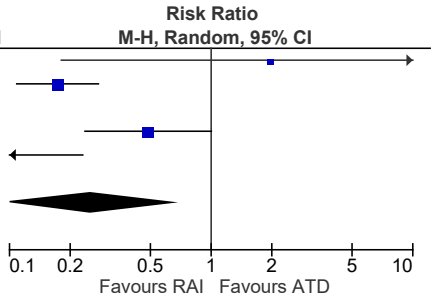


Figure 8: Osteoporosis

Study or Subgroup	RAI Events	RAI Total	ATD Events	ATD Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Torring 1996	6	34	5	36	100.0%	1.27 [0.43, 3.78]
Total (95% CI)		**34**		**36**	**100.0%**	**1.27 [0.43, 3.78]**
Total events	6		5			

Heterogeneity: Not applicable
Test for overall effect: Z = 0.43 (P = 0.67)

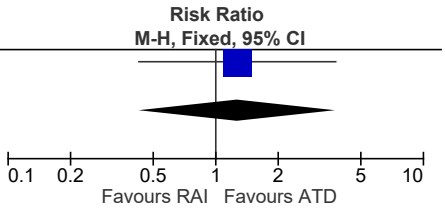


Figure 9: Agranulocytosis

Study or Subgroup	RAI Events	RAI Total	ATD Events	ATD Total	Weight	Peto Odds Ratio Peto, Fixed, 95% CI
Chen 2009	0	209	7	214	100.0%	0.13 [0.03, 0.60]
Total (95% CI)		**209**		**214**	**100.0%**	**0.13 [0.03, 0.60]**
Total events	0		7			

Heterogeneity: Not applicable
Test for overall effect: Z = 2.63 (P = 0.008)

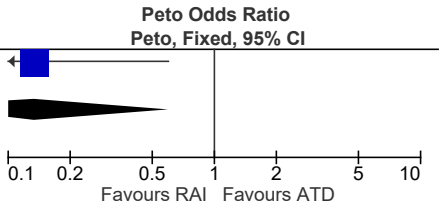


Figure 10: Severe liver damage

Study or Subgroup	RAI Events	RAI Total	ATD Events	ATD Total	Weight	Peto Odds Ratio Peto, Fixed, 95% CI
Chen 2009	0	209	5	214	100.0%	0.14 [0.02, 0.79]
Total (95% CI)		**209**		**214**	**100.0%**	**0.14 [0.02, 0.79]**
Total events	0		5			

Heterogeneity: Not applicable
Test for overall effect: Z = 2.22 (P = 0.03)

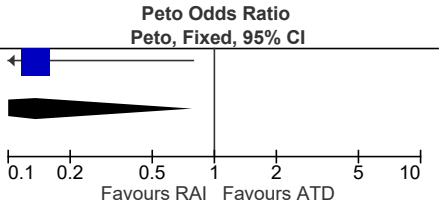


Figure 11: Malignancy

Study or Subgroup	RAI Events	RAI Total	ATD Events	ATD Total	Weight	Peto Odds Ratio Peto, Fixed, 95% CI
Chen 2009	0	209	0	177		Not estimable
Total (95% CI)		**209**		**177**		**Not estimable**
Total events	0		0			

Heterogeneity: Not applicable
Test for overall effect: Not applicable

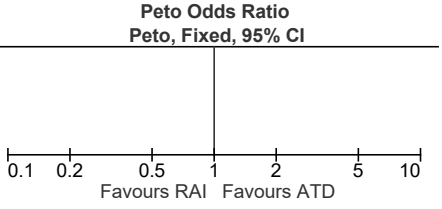
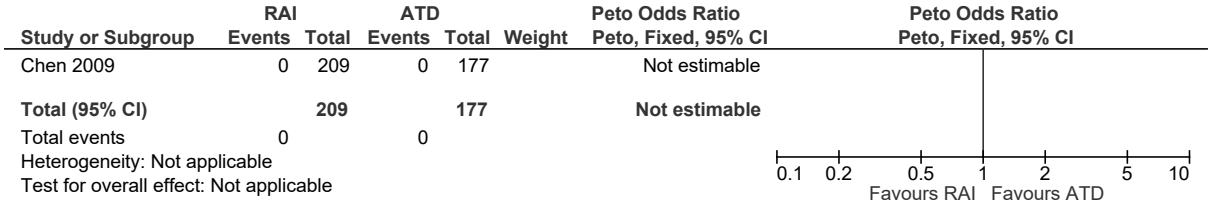


Figure 12: Thyroid storm



E.1.2 Surgery vs antithyroid drugs, adults with Graves' disease, first line treatment

Figure 13: Ophthalmopathy (new/worsening cases)

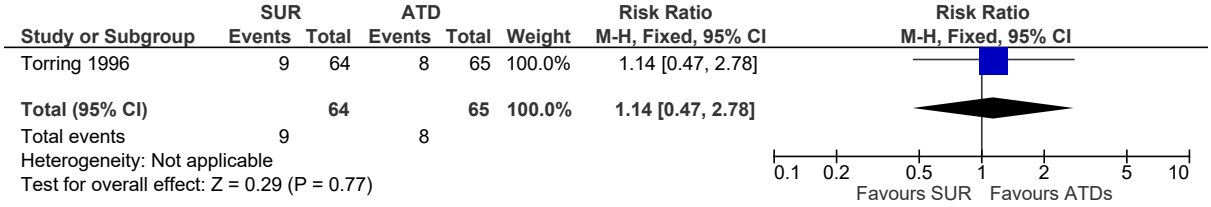


Figure 14: Osteoporosis

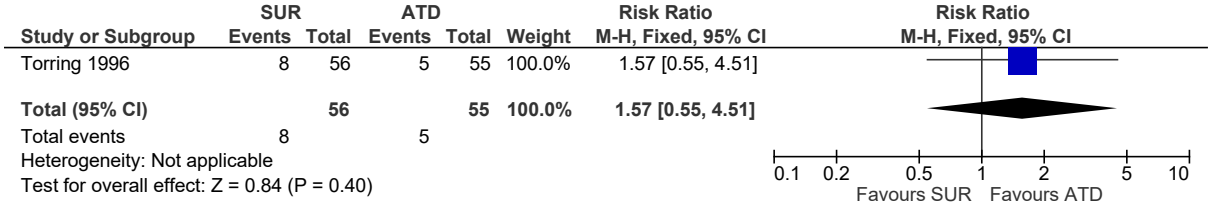
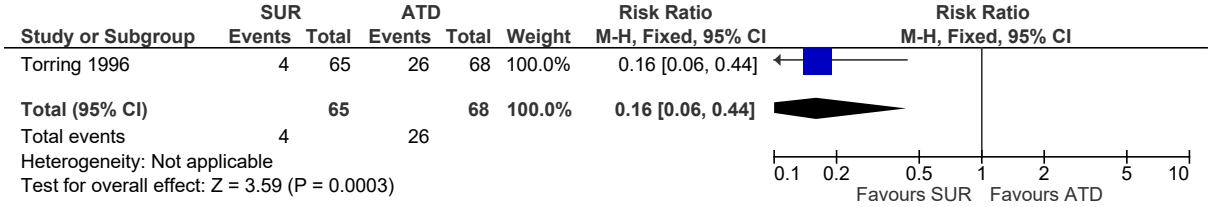


Figure 15: Hyperthyroidism



E.1.3 Radioactive iodine vs surgery, adults with Graves' disease, first line treatment

Figure 16: Ophthalmopathy (new/worsening cases)

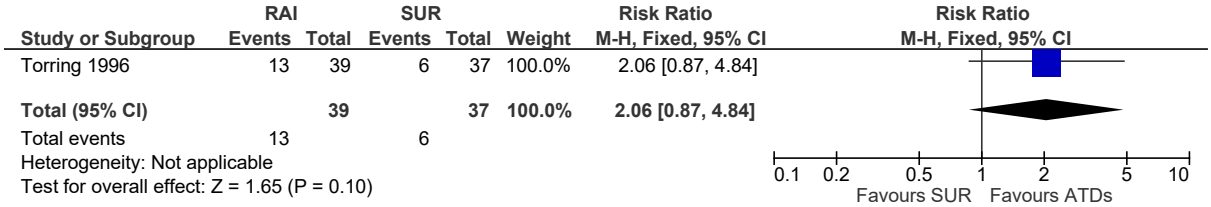


Figure 17: Osteoporosis

Study or Subgroup	RAI Events	RAI Total	SUR Events	SUR Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Torring 1996	6	34	7	34	100.0%	0.86 [0.32, 2.29]
Total (95% CI)		**34**		**34**	**100.0%**	**0.86 [0.32, 2.29]**
Total events	6		7			

Heterogeneity: Not applicable
Test for overall effect: Z = 0.31 (P = 0.76)

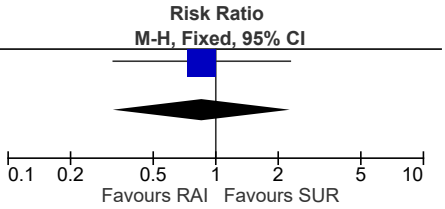
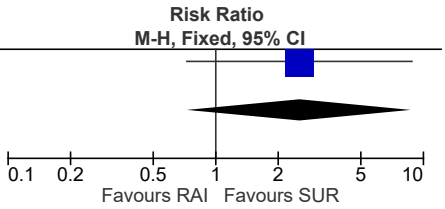


Figure 18: Hyperthyroidism

Study or Subgroup	RAI Events	RAI Total	SUR Events	SUR Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Torring 1996	8	39	3	37	100.0%	2.53 [0.73, 8.82]
Total (95% CI)		**39**		**37**	**100.0%**	**2.53 [0.73, 8.82]**
Total events	8		3			

Heterogeneity: Not applicable
Test for overall effect: Z = 1.46 (P = 0.15)



E.1.4 Radioactive iodine vs antithyroid drugs, adults with Graves' disease, second line treatment

Figure 19: Euthyroidism

Study or Subgroup	RAI Events	RAI Total	ATD Events	ATD Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Azizi 2005	16	41	26	28	100.0%	0.42 [0.28, 0.62]
Total (95% CI)		**41**		**28**	**100.0%**	**0.42 [0.28, 0.62]**
Total events	16		26			

Heterogeneity: Not applicable
Test for overall effect: Z = 4.29 (P < 0.0001)

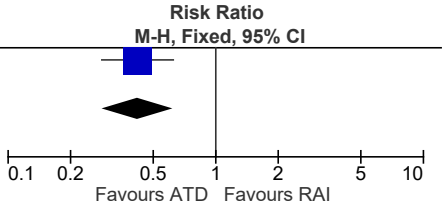


Figure 20: Hypothyroidism

Study or Subgroup	RAI Events	RAI Total	ATD Events	ATD Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Azizi 2005	25	41	1	28	100.0%	17.07 [2.45, 118.83]
Total (95% CI)		**41**		**28**	**100.0%**	**17.07 [2.45, 118.83]**
Total events	25		1			

Heterogeneity: Not applicable
Test for overall effect: Z = 2.87 (P = 0.004)

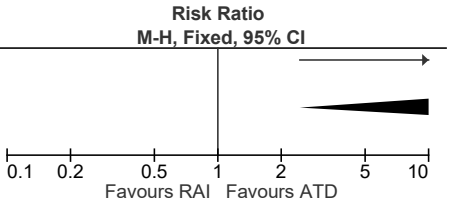


Figure 21: Hyperthyroidism

Study or Subgroup	RAI Events	RAI Total	ATD Events	ATD Total	Weight	Peto Odds Ratio Peto, Fixed, 95% CI
Azizi 2005	0	41	1	28	100.0%	0.09 [0.00, 4.60]
Total (95% CI)		**41**		**28**	**100.0%**	**0.09 [0.00, 4.60]**
Total events	0		1			

Heterogeneity: Not applicable
Test for overall effect: Z = 1.21 (P = 0.23)

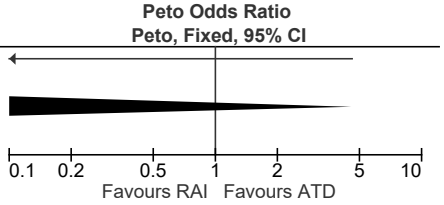
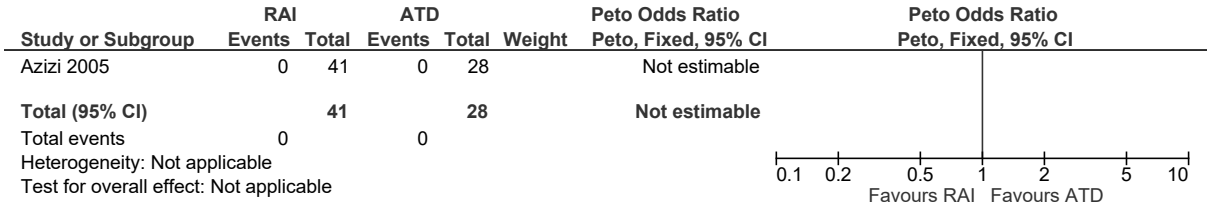


Figure 22: Agranulocytosis



E.2 Forest plots: Radioactive Iodine safety

E.2.1 Radioactive iodine vs surgery

Figure 23: Total cancer diagnoses (RR)

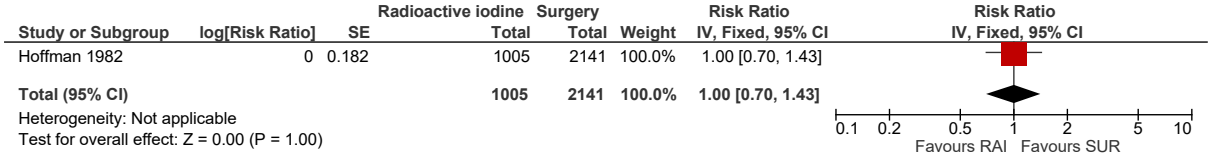


Figure 24: Total cancer diagnoses (HR)

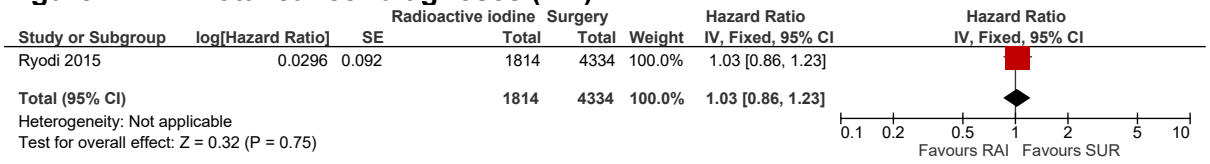


Figure 25: Total cancer mortality

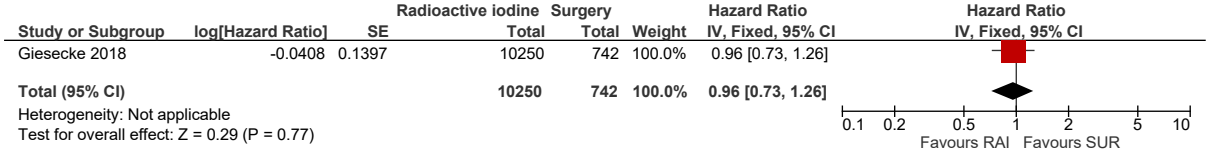


Figure 26: Lip, oral, pharynx cancer diagnoses

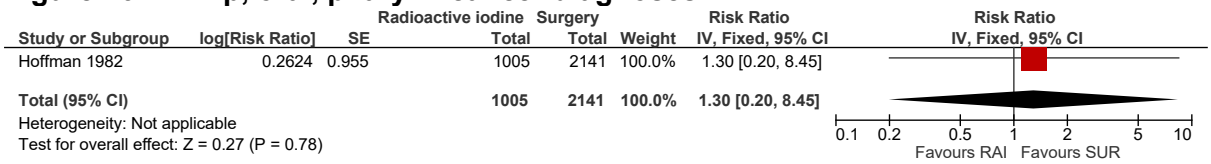


Figure 27: Digestive organ cancer diagnoses

Study or Subgroup	log[Risk Ratio]	SE	Radioactive iodine Total	Surgery Total	Weight	Risk Ratio IV, Fixed, 95% CI
Hoffman 1982	0.0953	0.3093	1005	2141	100.0%	1.10 [0.60, 2.02]
Total (95% CI)			**1005**	**2141**	**100.0%**	**1.10 [0.60, 2.02]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.31 (P = 0.76)

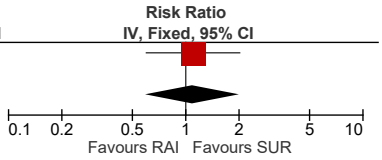


Figure 28: Respiratory cancer diagnoses

Study or Subgroup	log[Risk Ratio]	SE	Radioactive iodine Total	Surgery Total	Weight	Risk Ratio IV, Fixed, 95% CI
Hoffman 1982	0.2624	0.6014	1005	2141	100.0%	1.30 [0.40, 4.23]
Total (95% CI)			**1005**	**2141**	**100.0%**	**1.30 [0.40, 4.23]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.44 (P = 0.66)

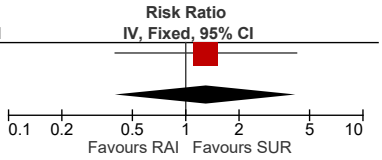


Figure 29: Breast cancer diagnoses

Study or Subgroup	log[Risk Ratio]	SE	Radioactive iodine Total	Surgery Total	Weight	Risk Ratio IV, Fixed, 95% CI
Hoffman 1982	-0.2231	0.2398	1005	2141	100.0%	0.80 [0.50, 1.28]
Total (95% CI)			**1005**	**2141**	**100.0%**	**0.80 [0.50, 1.28]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.93 (P = 0.35)

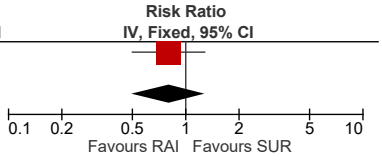


Figure 30: Genital cancer diagnoses

Study or Subgroup	log[Risk Ratio]	SE	Radioactive iodine Total	Surgery Total	Weight	Risk Ratio IV, Fixed, 95% CI
Hoffman 1982	0.0953	0.5161	1005	2141	100.0%	1.10 [0.40, 3.02]
Total (95% CI)			**1005**	**2141**	**100.0%**	**1.10 [0.40, 3.02]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.18 (P = 0.85)

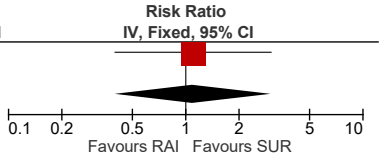


Figure 31: Kidney cancer diagnoses

Study or Subgroup	log[Risk Ratio]	SE	Radioactive iodine Total	Surgery Total	Weight	Risk Ratio IV, Fixed, 95% CI
Hoffman 1982	1.2238	0.978	1005	2141	100.0%	3.40 [0.50, 23.12]
Total (95% CI)			**1005**	**2141**	**100.0%**	**3.40 [0.50, 23.12]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.25 (P = 0.21)

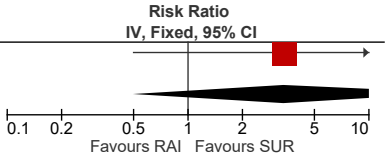


Figure 32: Melanoma diagnoses

Study or Subgroup	log[Risk Ratio]	SE	Radioactive iodine Total	Surgery Total	Weight	Risk Ratio IV, Fixed, 95% CI
Hoffman 1982	-5.8807	4.0485	1005	2141	100.0%	0.00 [0.00, 7.80]
Total (95% CI)			**1005**	**2141**	**100.0%**	**0.00 [0.00, 7.80]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.45 (P = 0.15)

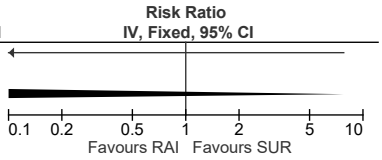


Figure 33: CNS cancer diagnoses

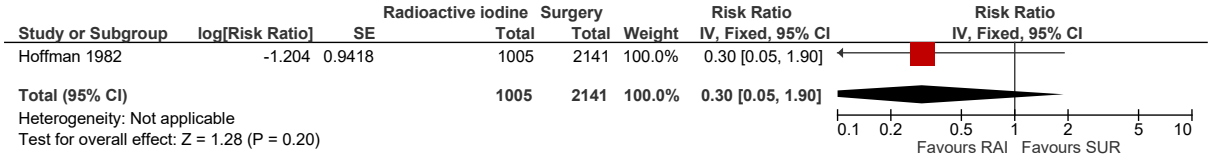


Figure 34: Thyroid cancer diagnoses

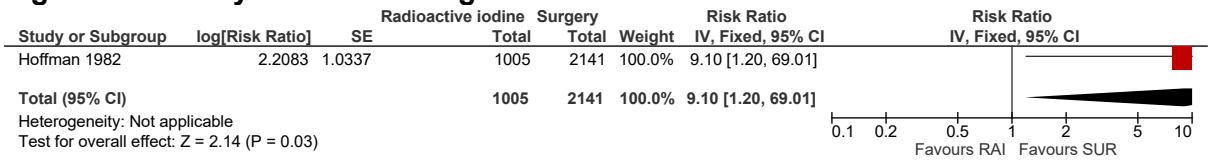


Figure 35: Other solid tumour diagnoses

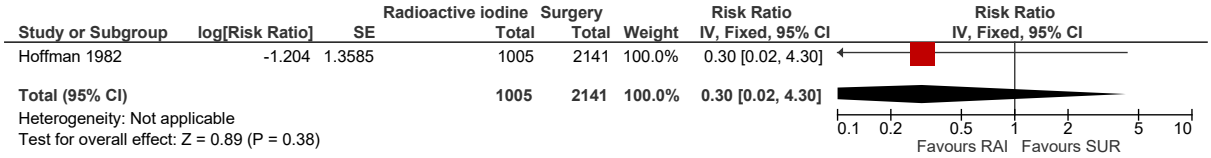


Figure 36: Lymphatic cancer diagnoses

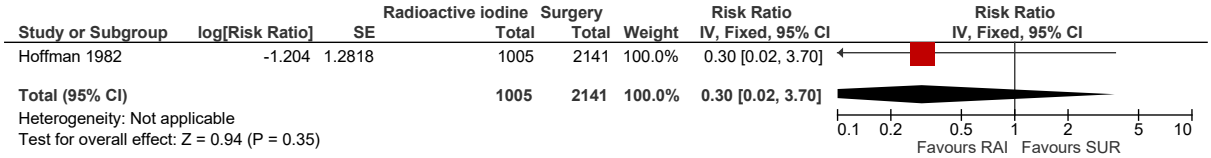
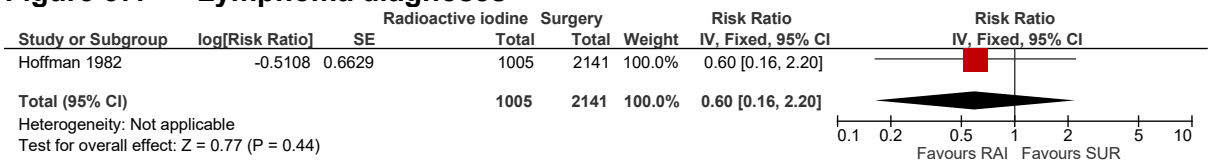


Figure 37: Lymphoma diagnoses



E.2.2 Radioactive iodine vs general population

Figure 38: Total cancer diagnoses

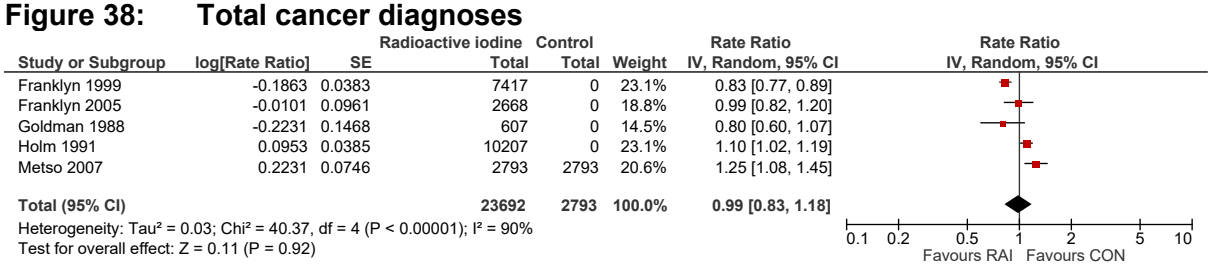


Figure 39: Lip, oral, pharynx cancer diagnoses

Study or Subgroup	log[Rate Ratio]	SE	Weight	Rate Ratio IV, Fixed, 95% CI
Franklyn 1999	-0.5978	0.4448	30.2%	0.55 [0.23, 1.32]
Holm 1991	0.207	0.3176	59.3%	1.23 [0.66, 2.29]
Metso 2007	-0.2357	0.7546	10.5%	0.79 [0.18, 3.47]
Total (95% CI)			**100.0%**	**0.92 [0.57, 1.49]**

Heterogeneity: Chi² = 2.21, df = 2 (P = 0.33); I² = 10%
Test for overall effect: Z = 0.34 (P = 0.74)

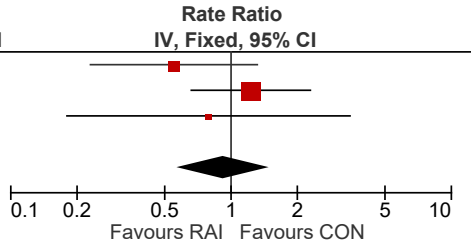


Figure 40: Salivary gland cancer diagnoses

Study or Subgroup	log[Rate Ratio]	SE	Weight	Rate Ratio IV, Fixed, 95% CI
Holm 1991	0.1398	1.0744	67.6%	1.15 [0.14, 9.45]
Metso 2007	1.6582	1.5534	32.4%	5.25 [0.25, 110.26]
Total (95% CI)			**100.0%**	**1.88 [0.33, 10.62]**

Heterogeneity: Chi² = 0.65, df = 1 (P = 0.42); I² = 0%
Test for overall effect: Z = 0.71 (P = 0.48)

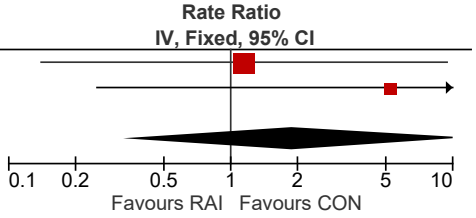


Figure 41: Digestive organs and peritoneum cancer diagnoses

Study or Subgroup	log[Rate Ratio]	SE	Weight	Rate Ratio IV, Random, 95% CI
Franklyn 1999	-0.1054	0.0796	35.4%	0.90 [0.77, 1.05]
Goldman 1988	-0.2231	0.3537	7.2%	0.80 [0.40, 1.60]
Holm 1991	0.1906	0.0973	32.1%	1.21 [1.00, 1.46]
Metso 2007	0.2151	0.1359	25.3%	1.24 [0.95, 1.62]
Total (95% CI)			**100.0%**	**1.06 [0.87, 1.30]**

Heterogeneity: Tau² = 0.02; Chi² = 7.95, df = 3 (P = 0.05); I² = 62%
Test for overall effect: Z = 0.60 (P = 0.55)

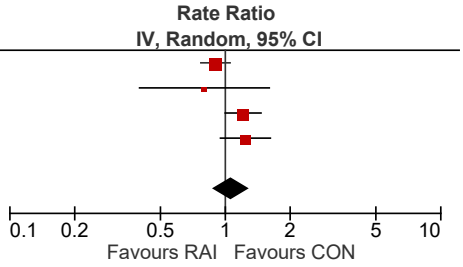


Figure 42: Bone, connective tissue and skin cancer diagnoses

Study or Subgroup	log[Rate Ratio]	SE	Weight	Rate Ratio IV, Fixed, 95% CI
Franklyn 1999	-0.1278	0.1468	78.4%	0.88 [0.66, 1.17]
Metso 2007	-0.1054	0.2799	21.6%	0.90 [0.52, 1.56]
Total (95% CI)			**100.0%**	**0.88 [0.69, 1.14]**

Heterogeneity: Chi² = 0.01, df = 1 (P = 0.94); I² = 0%
Test for overall effect: Z = 0.95 (P = 0.34)

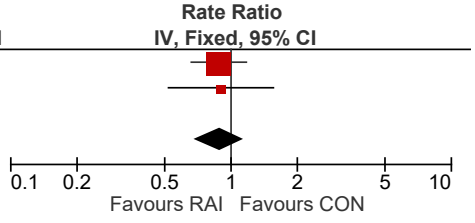


Figure 43: Breast cancer diagnoses

Study or Subgroup	log[Rate Ratio]	SE	Weight	Rate Ratio IV, Fixed, 95% CI
Franklyn 1999	0.0677	0.0883	44.1%	1.07 [0.90, 1.27]
Goldman 1988	0	0.2606	5.1%	1.00 [0.60, 1.67]
Holm 1991	0.0296	0.092	40.6%	1.03 [0.86, 1.23]
Metso 2007	0.4447	0.1829	10.3%	1.56 [1.09, 2.23]
Total (95% CI)			**100.0%**	**1.09 [0.97, 1.22]**

Heterogeneity: Chi² = 4.37, df = 3 (P = 0.22); I² = 31%
Test for overall effect: Z = 1.49 (P = 0.14)

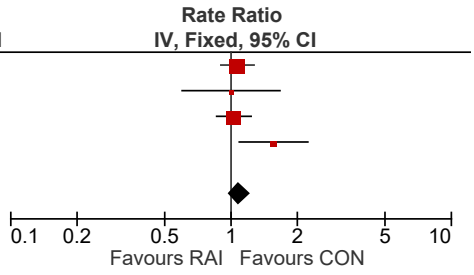


Figure 44: Brain and other CNS cancer diagnoses

Study or Subgroup	log[Rate Ratio]	SE	Weight	Rate Ratio IV, Fixed, 95% CI
Franklyn 1999	-0.6162	0.5068	12.1%	0.54 [0.20, 1.46]
Goldman 1988	0.9555	1.1018	2.6%	2.60 [0.30, 22.53]
Holm 1991	0.4886	0.2007	77.3%	1.63 [1.10, 2.42]
Metso 2007	0.6098	0.6255	8.0%	1.84 [0.54, 6.27]
Total (95% CI)			**100.0%**	**1.46 [1.03, 2.06]**

Heterogeneity: Chi² = 4.56, df = 3 (P = 0.21); I² = 34%
Test for overall effect: Z = 2.13 (P = 0.03)

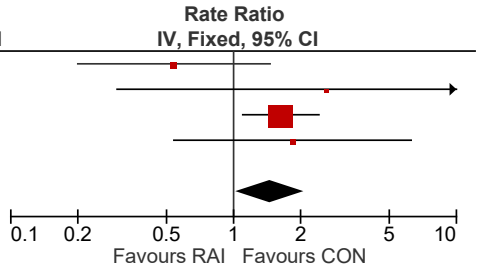


Figure 45: Respiratory cancer diagnoses

Study or Subgroup	log[Rate Ratio]	SE	Weight	Rate Ratio IV, Random, 95% CI
Franklyn 1999	-0.5108	0.1139	37.6%	0.60 [0.48, 0.75]
Holm 1991	0.157	0.1512	35.4%	1.17 [0.87, 1.57]
Metso 2007	-0.1393	0.2725	27.0%	0.87 [0.51, 1.48]
Total (95% CI)			**100.0%**	**0.84 [0.52, 1.35]**

Heterogeneity: Tau² = 0.14; Chi² = 12.65, df = 2 (P = 0.002); I² = 84%
Test for overall effect: Z = 0.72 (P = 0.47)

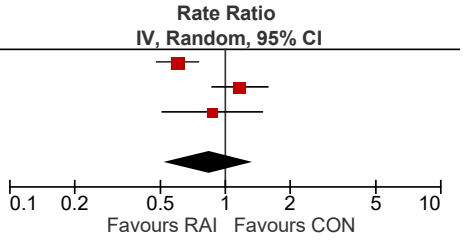


Figure 46: Genitourinary cancer diagnoses

Study or Subgroup	log[Rate Ratio]	SE	Weight	Rate Ratio IV, Random, 95% CI
Franklyn 1999	-0.2744	0.1039	39.0%	0.76 [0.62, 0.93]
Holm 1991	0.0677	0.1115	37.7%	1.07 [0.86, 1.33]
Metso 2007	0.131	0.2069	23.3%	1.14 [0.76, 1.71]
Total (95% CI)			**100.0%**	**0.95 [0.73, 1.24]**

Heterogeneity: Tau² = 0.04; Chi² = 6.29, df = 2 (P = 0.04); I² = 68%
Test for overall effect: Z = 0.37 (P = 0.71)

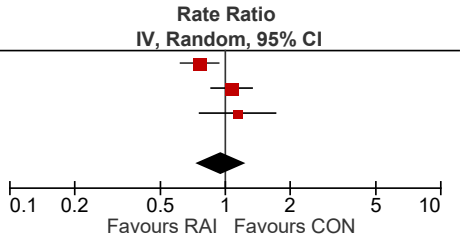


Figure 47: Thyroid cancer diagnoses

Study or Subgroup	log[Rate Ratio]	SE	Weight	Rate Ratio IV, Fixed, 95% CI
Franklyn 1999	1.1787	0.3336	51.9%	3.25 [1.69, 6.25]
Holm 1991	0.2776	0.3938	37.2%	1.32 [0.61, 2.86]
Metso 2007	0.5596	0.7281	10.9%	1.75 [0.42, 7.29]
Total (95% CI)			**100.0%**	**2.17 [1.36, 3.48]**

Heterogeneity: Chi² = 3.15, df = 2 (P = 0.21); I² = 36%
Test for overall effect: Z = 3.23 (P = 0.001)

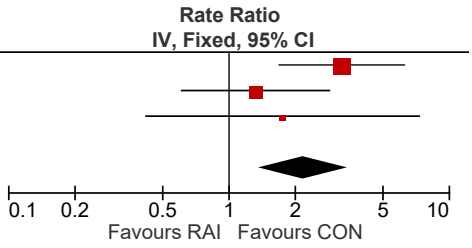


Figure 48: Haematopoietic cancer diagnoses

Study or Subgroup	log[Rate Ratio]	SE	Weight	Rate Ratio IV, Random, 95% CI
Franklyn 1999	-0.4943	0.2027	38.7%	0.61 [0.41, 0.91]
Holm 1991	-0.1985	0.2131	37.1%	0.82 [0.54, 1.25]
Metso 2007	0.2469	0.3153	24.2%	1.28 [0.69, 2.37]
Total (95% CI)			**100.0%**	**0.81 [0.56, 1.19]**

Heterogeneity: Tau² = 0.06; Chi² = 3.99, df = 2 (P = 0.14); I² = 50%
Test for overall effect: Z = 1.06 (P = 0.29)

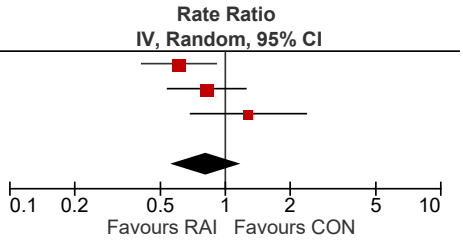


Figure 49: Kidney cancer diagnoses

Study or Subgroup	log[Rate Ratio]	SE	Weight	Rate Ratio IV, Fixed, 95% CI
Holm 1991	0.4121	0.1805	83.4%	1.51 [1.06, 2.15]
Metso 2007	0.8502	0.404	16.6%	2.34 [1.06, 5.17]
Total (95% CI)			**100.0%**	**1.62 [1.18, 2.24]**

Heterogeneity: Chi² = 0.98, df = 1 (P = 0.32); I² = 0%
Test for overall effect: Z = 2.94 (P = 0.003)

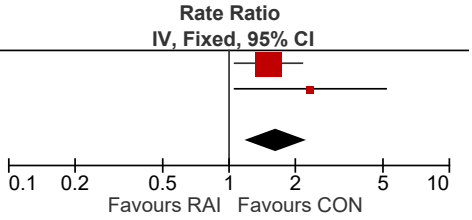


Figure 50: Parathyroid cancer diagnoses

Study or Subgroup	log[Rate Ratio]	SE	Weight	Rate Ratio IV, Fixed, 95% CI
Holm 1991	0.47	0.2936	100.0%	1.60 [0.90, 2.84]
Total (95% CI)			**100.0%**	**1.60 [0.90, 2.84]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.60 (P = 0.11)

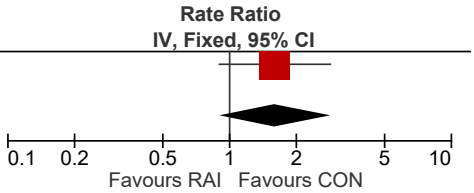
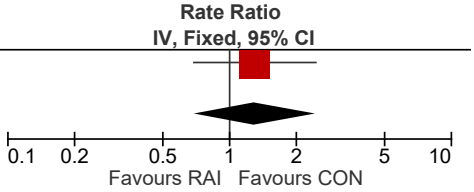


Figure 51: Prostate cancer diagnoses

Study or Subgroup	log[Rate Ratio]	SE	Weight	Rate Ratio IV, Fixed, 95% CI
Metso 2007	0.2624	0.3232	100.0%	1.30 [0.69, 2.45]
Total (95% CI)			**100.0%**	**1.30 [0.69, 2.45]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.81 (P = 0.42)



Appendix F: GRADE tables

F.1 Drugs vs Surgery vs Radioactive Iodine

Table 18: Clinical evidence profile: Radioactive iodine vs antithyroid drugs, Graves' disease, first line treatment

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RAI	ATD	Relative (95% CI)	Absolute		
Mortality (follow-up 9 years)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	0/209 (0%)	0%	-	not estimable [5]	⊕⊕OO LOW	CRITICAL
Ophthalmopathy (new/worsening cases) (follow-up 1-9 years)												
4	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	125/487 (25.7%)	10.3%	RR 2.17 (1.64 to 2.88)	121 more per 1000 (from 66 more to 194 more)	⊕⊕⊕O MODERATE	IMPORTANT
Euthyroidism (follow-up 1-9 years; assessed with: (at end of follow-up))												
3	randomised trials	serious[1]	very serious[3]	no serious indirectness	no serious imprecision	none	278/387 (71.8%)	75.9%	RR 0.78 (0.37 to 1.62)	167 fewer per 1000 (from 478 fewer to 471 more)	⊕OOO VERY LOW	IMPORTANT
Hypothyroidism (follow-up 1-9 years; assessed with: (at end of follow-up))												
3	randomised	serious[1]	no serious	no serious	no serious	none	63/387	3.4%	RR 5.89 (3.12	166 more per 1000 (from	⊕⊕⊕O	IMPORTANT

	trials		inconsistency	indirectness	imprecision		(16.3%)		to 11.11)	72 more to 344 more)	MODERATE	
Hyperthyroidism (persistence/recurrence) (follow-up 1-9 years)												
5	randomised trials	serious[1]	serious[3]	no serious indirectness	no serious imprecision	none	30/573 (5.2%)	24.1%	RR 0.25 (0.09 to 0.69)	181 fewer per 1000 (from 75 fewer to 219 fewer)	⊕⊕OO LOW	IMPORTANT
Osteoporosis (follow-up 3 years)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[4]	none	6/34 (17.6%)	13.9%	RR 1.27 (0.43 to 3.78)	38 more per 1000 (from 79 fewer to 386 more)	⊕OOO VERY LOW	IMPORTANT
Agranulocytosis (follow-up 9 years)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	0/209 (0%)	3.3%	Peto OR 0.13 (0.03 to 0.6)	29 fewer per 1000 (from 13 fewer to 32 fewer)	⊕⊕OO LOW	IMPORTANT
Severe liver damage (follow-up 9 years)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[4]	none	0/209 (0%)	2.3%	OR 0.14 (0.02 to 0.79)	20 fewer per 1000 (from 5 fewer to 23 fewer)	⊕⊕OO LOW	IMPORTANT
Malignancy (follow-up 9 years)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	0/209 (0%)	0%	-	not estimable [5]	⊕⊕OO LOW	IMPORTANT
Thyroid storm (follow-up 9 years)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	0/209 (0%)	0%	-	not estimable [5]	⊕⊕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 increment as zero events in at least one arm
[3] Downgraded by 1 or 2 increments because the point estimate and or the confidence intervals varied widely across studies, unexplained by subgroup analysis
[4] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs
[5] Zero events in both arms

Table 19: Clinical evidence profile: Surgery vs antithyroid drugs, Graves' disease, first line treatment

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	SUR	ATD	Relative (95% CI)	Absolute		
Ophthalmopathy (follow-up 4 years; assessed with: (new/worsening cases))												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	none	9/64 (14.1%)	12.3%	RR 1.14 (0.47 to 2.78)	17 more per 1000 (from 65 fewer to 219 more)	⊕⊕OO LOW	IMPORTANT
Osteoporosis (follow-up 14-21 years)												
1	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	very serious[1]	none	8/56 (14.3%)	9.1%	RR 1.57 (0.55 to 4.51)	52 more per 1000 (from 41 fewer to 319 more)	⊕OOO VERY LOW	IMPORTANT
Hyperthyroidism (follow-up 4 years; assessed with: (persistence/recurrence))												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	4/65 (6.2%)	38.2%	RR 0.16 (0.06 to 0.44)	321 fewer per 1000 (from 214 fewer to 359 fewer)	⊕⊕⊕⊕ HIGH	IMPORTANT

[1] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs [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 20: Clinical evidence profile: Radioactive iodine vs surgery, Graves' disease, first line treatment

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RAI	Control	Relative (95% CI)	Absolute		
Ophthalmopathy (follow-up 4 years; assessed with: (new/worsening cases))												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	13/39 (33.3%)	16.2%	RR 2.06 (0.87 to 4.84)	172 more per 1000 (from 21 fewer to 622 more)	⊕⊕⊕O MODERATE	IMPORTANT
Osteoporosis (follow-up 14-21 years)												
1	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	very serious[1]	none	6/34 (17.6%)	20.6%	RR 0.86 (0.32 to 2.29)	29 fewer per 1000 (from 140 fewer to 266 more)	⊕OOO VERY LOW	IMPORTANT
Hyperthyroidism (persistence/recurrence) (follow-up 4 years)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[1]	none	8/39 (20.5%)	8.1%	RR 2.53 (0.73 to 8.82)	124 more per 1000 (from 22 fewer to 633 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 21: Clinical evidence profile: Radioactive iodine vs antithyroid drugs, Graves' disease, second line treatment

Quality assessment							No of patients		Effect		Quality	Importance
No of	Design	Risk of	Inconsistency	Indirectness	Imprecision	Other	RAI	ATD	Relative	Absolute		

studies		bias				considerations			(95% CI)			
Euthyroidism (follow-up 10 years; assessed with: (at end of follow-up))												
1	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	16/41 (39%)	92.9%	RR 0.42 (0.28 to 0.62)	539 fewer per 1000 (from 353 fewer to 669 fewer)	⊕⊕OO LOW	IMPORTANT
Hypothyroidism (follow-up 10 years; assessed with: (at end of follow-up))												
1	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	25/41 (61%)	3.6%	RR 17.07 (2.45 to 118.83)	579 more per 1000 (from 52 more to 1000 more)	⊕⊕OO LOW	IMPORTANT
Hyperthyroidism (follow-up 10 years; assessed with: (at end of follow-up))												
1	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	very serious[1]	none	0/41 (0%)	3.6%	OR 0.09 (0 to 4.6)	33 fewer per 1000 (from 36 fewer to 111 more)	⊕OOO VERY LOW	IMPORTANT
Agranulocytosis (follow-up 10 years)												
1	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	0/41 (0%)	0%	-	not estimable[3]	⊕OOO VERY LOW	IMPORTANT

[1] Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
[2] Downgraded by 1 increment as at least one arm with zero events
[3] zero events in both arms

F.2 Radioactive iodine safety

Table 22: Clinical evidence profile: radioactive iodine vs surgery

Quality assessment						No of patients	Effect	Quality	Importance

No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Radioactive iodine**	**Surgery**	**Relative (95% CI)**	**Absolute**		
Total cancer diagnoses (RR) (follow-up median 15 years)												
1	observational studies	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	1005	11.5%	RR 1.00 (0.7 to 1.43)	0 fewer per 1000 (from 35 fewer to 49 more)	⊕OOO VERY LOW	CRITICAL
Total cancer diagnoses (HR) (follow-up median 10 years)												
1	observational studies	very serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	1814	0%[3]	HR 1.02 (0.86 to 1.23)	Not estimable	⊕OOO VERY LOW	CRITICAL
Total cancer mortality (follow-up median 16.3 years)												
1	observational studies	no serious risk of bias[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	10250	0%[3]	HR 0.96 (0.73 to 1.26)	Not estimable	⊕OOO VERY LOW	CRITICAL
Lip, oral, pharynx cancer diagnoses (follow-up median 15 years)												
1	observational studies	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	1005	0.37%	RR 1.3 (0.2 to 8.45)	1 more per 1000 (from 3 fewer to 28 more)	⊕OOO VERY LOW	CRITICAL
Digestive organ and peritoneum cancer diagnoses (follow-up median 15 years)												
1	observational studies	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	1005	2.38%	RR 1.1 (0.6 to 2.02)	2 more per 1000 (from 10 fewer to 24 more)	⊕OOO VERY LOW	CRITICAL
Respiratory cancer diagnoses (follow-up median 15 years)												
1	observational studies	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	1005	0.7%	RR 1.3 (0.4 to 4.23)	2 more per 1000 (from 4 fewer to 23 more)	⊕OOO VERY LOW	CRITICAL
Breast cancer diagnoses (follow-up median 15 years)												

1	observational studies	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	1005	3.41%	RR 0.8 (0.5 to 1.28)	7 fewer per 1000 (from 17 fewer to 10 more)	⊕OOO VERY LOW	CRITICAL
Genital cancer diagnoses (follow-up 15 years)												
1	observational studies	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	1005	2.1%	RR 1.1 (0.4 to 3.02)	2 more per 1000 (from 13 fewer to 42 more)	⊕OOO VERY LOW	CRITICAL
Kidney and bladder cancer diagnoses (follow-up median 15 years)												
1	observational studies	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	1005	0.19%	RR 3.4 (0.5 to 23.12)	5 more per 1000 (from 1 fewer to 42 more)	⊕OOO VERY LOW	CRITICAL
Melanoma cancer diagnoses (follow-up median 15 years)												
1	observational studies	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	1005	0.05%	RR 0 (0 to 7.8)	1 fewer per 1000 (from 1 fewer to 3 more)	⊕OOO VERY LOW	CRITICAL
CNS cancer diagnoses (follow-up median 15 years)												
1	observational studies	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	1005	0.28%	RR 0.3 (0.05 to 1.9)	2 fewer per 1000 (from 3 fewer to 3 more)	⊕OOO VERY LOW	CRITICAL
Thyroid cancer diagnoses (follow-up median 15 years)												
1	observational studies	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	1005	0.05%	RR 9.1 (1.2 to 69.01)	4 more per 1000 (from 0 more to 34 more)	⊕OOO VERY LOW	CRITICAL
Other solid tumour cancer diagnoses (follow-up median 15 years)												
1	observational studies	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	1005	0.28%	RR 0.3 (0.02 to 4.3)	2 fewer per 1000 (from 3 fewer to 9 more)	⊕OOO VERY LOW	CRITICAL
Lymphatic cancer diagnoses (follow-up median 15 years)												
1	observational	serious[1]	no serious	no serious	very serious[2]	none	1005	0.33%	RR 0.3 (0.02	2 fewer per 1000	⊕OOO	CRITICAL

	studies		inconsistency	indirectness					to 3.7)	(from 3 fewer to 9 more)	VERY LOW	
Leukaemia diagnoses (follow-up median 15 years)												
1	observational studies	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	1005	0.47%	RR 0.6 (0.16 to 2.2)	2 fewer per 1000 (from 4 fewer to 6 more)	⊕OOO VERY LOW	CRITICAL

[1] Default starting quality of low overall due to selection bias in non-randomised studies. Downgraded further for risk of bias if the majority of evidence was at additional risk of bias, either once if high risk of bias or twice if very high risk of bias
[2] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs
[3] No control group risk provided

Table 23: Clinical evidence profile: radioactive iodine vs general population

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Radioactive iodine**	**General population**	**Relative (95% CI)**	**Absolute**		
Total cancer diagnoses (follow-up 5-17 years)												
5	observational studies	serious[1]	serious[2]	no serious indirectness	no serious imprecision	none	23692	7.4%	RR 0.99 (0.83 to 1.18)	1 fewer per 1000 (from 13 fewer to 13 more)	⊕OOO VERY LOW	CRITICAL
Lip, oral, pharynx cancer diagnoses (follow-up 5-15 years)												
3	observational studies	serious[1]	no serious inconsistency	no serious indirectness	very serious[3]	none	20417	0.1%	RR 0.92 (0.57 to 1.49)	0 fewer per 1000 (from 0 fewer to 0 more)	⊕OOO VERY LOW	CRITICAL
Salivary gland cancer diagnoses (follow-up 10-15 years)												
2	observational studies	serious[1]	no serious inconsistency	no serious indirectness	very serious[3]	none	13000	0.01%	RR 1.88 (0.33 to 10.62)	0 more per 1000 (from 0 fewer to 1 more)	⊕OOO VERY LOW	CRITICAL

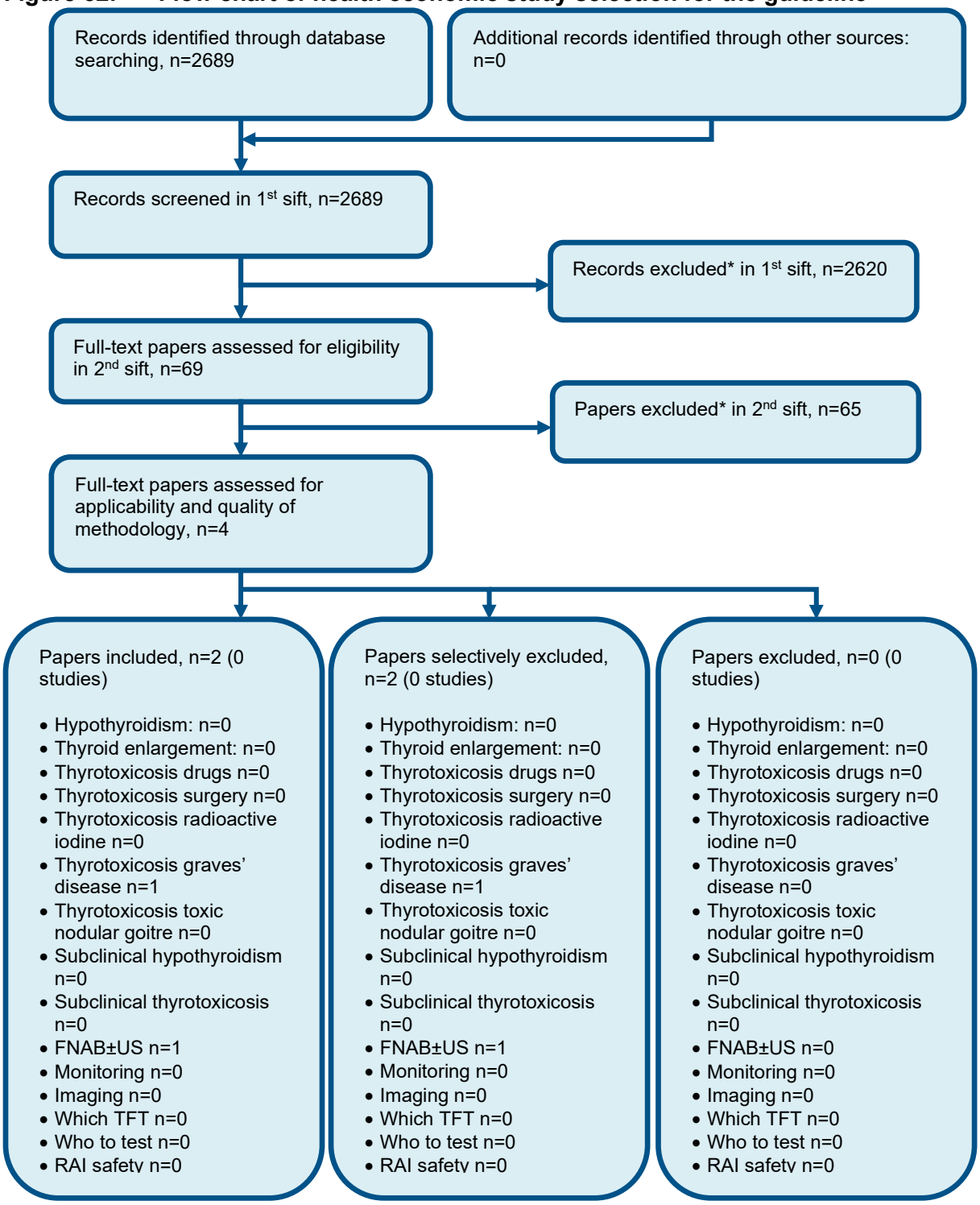
Digestive organs and peritoneum cancer diagnoses (follow-up 5-17 years)												
4	observational studies	serious[1]	serious[2]	no serious indirectness	no serious imprecision	none	21024	2.7%	RR 1.06(0.87 to 1.30)	2 more per 1000 (from 4 fewer to 8 more)	⊕OOO VERY LOW	CRITICAL
Bone, connective tissue and skin cancer diagnoses (follow-up 5-10 years)												
2	observational studies	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	10210	1.3%	RR 0.88 (0.69 to 1.14)	2 fewer per 1000 (from 4 fewer to 2 more)	⊕OOO VERY LOW	CRITICAL
Breast cancer diagnoses (follow-up 5-17 years)												
4	observational studies	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	21024	1.7%	RR 1.09 (0.97 to 1.22)	2 more per 1000 (from 1 fewer to 4 more)	⊕OOO VERY LOW	CRITICAL
Brain and other CNS cancer diagnoses (follow-up 5-17 years)												
4	observational studies	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	21024	0.3%	RR 1.46 (1.03 to 2.06)	1 more per 1000 (from 0 more to 3 more)	⊕OOO VERY LOW	CRITICAL
Respiratory cancer diagnoses (follow-up 5-17 years)												
3	observational studies	serious[1]	serious[2]	no serious indirectness	serious[3]	none	20417	0.9%	RR 0.84 (0.52 to 1.35)	1 fewer per 1000 (from4 fewer to 3 more)	⊕OOO VERY LOW	CRITICAL
Genitourinary cancer diagnoses (follow-up 5-17 years)												
3	observational studies	serious[1]	serious[2]	no serious indirectness	serious[3]	none	20417	1.6%	RR 0.95 (0.73 to 1.24)	1 fewer per 1000 (from 4 fewer to 4 more)	⊕OOO VERY LOW	CRITICAL
Thyroid cancer diagnoses (follow-up 5-17 years)												
3	observational studies	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	20417	0.1%	RR 2.17 (1.36 to 3.48)	1 more per 1000 (from 0 more to 2 more)	⊕OOO VERY LOW	CRITICAL
Haematopoietic cancer diagnoses (follow-up 5-17 years)												

3	observational studies	serious[1]	serious[2]	no serious indirectness	serious[3]	none	20417	0.5%	RR 0. 81 (0.56 to 1.19)	1 fewer per 1000 (from 2 fewer to 0 more)	⊕OOO VERY LOW	CRITICAL
Kidney cancer diagnoses (follow-up 10-15 years)												
2	observational studies	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	13000	0.4%	RR 1.62 (1.18 to 2.24)	2 more per 1000 (from 1 more to 5 more)	⊕OOO VERY LOW	CRITICAL
Parathyroid cancer diagnoses (follow-up 15 years)												
1	observational studies	serious[1]	no serious inconsistency	no serious indirectness	serious[3]	none	10207	0.22%	RR 1.6 (0.9 to 2.84)	1 more per 1000 (from 0 fewer to 4 more)	⊕OOO VERY LOW	CRITICAL
Prostate cancer diagnoses (follow-up 10 years)												
1	observational studies	serious[1]	no serious inconsistency	no serious indirectness	very serious[3]	none	2793	3.7%	RR 1.3 (0.69 to 2.45)	11 more per 1000 (from 11 fewer to 54 more)	⊕OOO VERY LOW	CRITICAL

[1] Default starting quality of low overall due to selection bias in non-randomised studies. Downgraded further for risk of bias if the majority of evidence was at additional risk of bias, either once if high risk of bias or twice if 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

Appendix G: Health economic evidence selection

Figure 52: 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

H.1 Drugs vs Surgery vs Radioactive iodine

Study	Donovan et al, 2016[35]			
Study details	**Population & interventions**	**Costs**	**Health outcomes**	**Cost effectiveness**
Economic analysis: CUA (health outcome: QALYs) **Study design:** Deterministic decision analytic model **Approach to analysis:** Markov model, cyclical and tracks key clinical options and outcomes of persons with Graves' disease following each of the 3 interventions. 3-monthly cycles. **Perspective:** UK NHS **Time Horizon:** lifetime **Treatment effect duration:** NR **Discounting:** Costs: 3.5% ; Outcomes: 3.5%	**Population:** People diagnosed with Graves' disease. **Cohort settings:** Start age: 40 years old women **Intervention 1:** Radioactive iodine (RAI) **Intervention 2:** Antithyroid drug (ATD) (carbimazole 5mg). **Intervention 3:** Total thyroidectomy (TT).	**Total costs (mean per patient):** Intervention 1: £5,425 Intervention 2: £16,866 Intervention 3: £7,115 Incremental (2−1):£11,441 (95% CI: NR; p=NR) Incremental (3−1): £1,690 (95% CI: NR; p=NR) Incremental (3−2): saves £9,751 (95% CI: NR; p=NR) **Currency & cost year:** 2015 UK pounds **Cost components incorporated:** • Long-term costs of medications • medical practitioner visits • pathology tests	**QALYs (mean per patient):** Intervention 1: 34.73 Intervention 2: 35.17 Intervention 3: 33.93 Incremental (2−1): 0.44 (95% CI: NR; p=NR) Incremental (3−1): −0.8 (95% CI: NR; p=NR) Incremental (3−2): −1.24 (95% CI: NR; p=NR)	**Full incremental analysis:** RAI dominated TT. At cost effectiveness threshold of £20,000 per QALY gained, RAI is cost-effective compared to ATD, while at a cost effectiveness threshold of £30,000 per QALY-gained; ATD is the cost-effective alternative, (ICER £26,279 per QALY-gained) compared to RAI. **Analysis of uncertainty:** One-way sensitivity analyses, where the value of a single parameter is changed across range of values with ICER values calculated. Transition probabilities ranges were based on 95% CI from published literature. Costs were varied from 50% to 150% depending on base case values. Results from these sensitivity analyses showed that ATD was a cost-effective alternative to RAI in most sensitivity analyses (calculated ICER remained below the £30,000 threshold). RAI was dominant over TT in all sensitivity

		associated with treatments and their complications.		analyses of all parameters assessed.
Data sources				
Health outcomes: Effectiveness data for the three interventions were based on a literature review that identified rates of efficacy, relapse, complications and HRQoL values associated with each treatment option. Some assumptions were made e.g. the failure rate with ATD and the incidence of hypothyroid post third dose of RAI. **Quality-of-life weights:** Effectiveness was evaluated by using the HRQoL estimates (health utilities) from published data, Euro-Qol – 5 dimensions or SF-36 values mapped to EQ-5D. Some of the values were also based on expert opinion using Delphi methodology. **Cost sources:** Unit costs based on 2015 UK sources (BNF, National Tariff). Where unit costs were not available, estimates were obtained from published literature or currency conversion.				
Comments				
Source of funding: An NHMRC early career fellowship (APPP1092153) support. **Limitations:** The estimates of relative treatment effects are not based on met-analysis of all the available evidence. Some costs have been based on the national tariff Payment System and maybe overestimated. The model has not been run probabilistically, to adequately assess parameter uncertainty.				
Overall applicability:[c] Directly applicable		**Overall quality:**[d] Minor limitations		

Abbreviations: CI: 95% confidence interval; CUA: cost–utility analysis; EQ-5D: Euro-qol 5 dimensions (scale: 0.0 [death] to 1.0 [full health], negative values mean worse than death); ICER: incremental cost-effectiveness ratio; NR: not reported; QALYs: quality-adjusted life years

(a) Directly applicable / Partially applicable / Not applicable

(b) Minor limitations / Potentially serious limitations / Very serious limitations

H.2 Radioactive iodine safety

None

Appendix I: Health economic analysis

I.1 Drugs vs Surgery vs Radioactive Iodine

None

I.2 Radioactive iodine safety

None

Appendix J: Excluded studies

J.1 Excluded clinical studies

J.1.1 Drugs vs Surgery vs Radioactive Iodine

Table 24: Studies excluded from the clinical review

Study	Exclusion reason
Abraham 2010[4]	Systematic review is not relevant to review question or unclear PICO
Abraham-nordling 2007[2]	No usable outcomes
Allannic 1990[5]	Incorrect interventions
Andrade 1999[6]	Less than minimum duration
Andrade 2001[7]	Incorrect interventions
Andrade 2004[8]	Incorrect interventions
Azizi 2012[12]	Wrong study design
Azizi 2018 [10]	NRS where RCTs are available
Barczynski 2012[13]	Incorrect interventions
Barczynski 2010[14]	Abstract only
Barczynski 2018[15]	Incorrect interventions
Benker 1995[18]	Incorrect interventions
Benker 1998[17]	Incorrect interventions
Bonnema 2003[19]	Incorrect interventions
Bonnema 2004[20]	Incorrect interventions
Bonnema 2011[21]	Inappropriate comparison
Braga 2002[22]	Less than minimum duration
Burch 2001[23]	No usable outcomes
Buscemi 2007[24]	Not guideline condition
Canto 2016[25]	Incorrect interventions
Chen 2011[29]	Inappropriate comparison
Chen 2014[30]	No additional outcomes to those reported elsewhere
Chi 2005[31]	Inappropriate comparison
Connell 1987[32]	No usable outcomes
De Luca 2018[33]	SR, checked for references
Edmonds 1994[36]	Incorrect interventions
Esfahani 2005[37]	Inappropriate comparison
García-mayor 1992[41]	Incorrect interventions
Glinoer 2001[43]	Incorrect interventions
Goni iriarte 1995[45]	Not in English
Grebe 1998[46]	Incorrect interventions
Hamide 2014[52]	NRS where RCTs are available
Hashizume 1991[53]	NRS without adequate adjustment
He 2004[54]	Incorrect interventions
Hoermann 2002[56]	Incorrect interventions
Homsanit 2001[63]	Incorrect interventions
Howarth 2001[64]	Incorrect interventions

Study	Exclusion reason
Jaiswal 2014[65]	Incorrect interventions
Järhult 2005[66]	Incorrect interventions
Jorde 1995[67]	Incorrect interventions
Kallner 1996[69]	Incorrect interventions
Kung 1995[71]	Incorrect interventions
Leclere 1994[72]	Not in English
Leslie 2003[73]	Incorrect interventions
Leung 2017 [74]	SR, checked for references
Li 2016[75]	SR, checked for references
Liu 2015[77]	Incorrect interventions
Liu 2017[76]	Incorrect interventions
Ljunggren 1998[78]	No usable outcomes
Lucas 1997[79]	Incorrect interventions
Ma 2008[80]	SR, checked for references
Ma 2016[81]	SR checked for references
Marcocci 1989[82]	Incorrect interventions
Mashio 1997[83]	Inappropriate comparison
Mastorakos 2003[84]	Incorrect interventions
Maugendre 1999[85]	Incorrect interventions
Mciver 1996[86]	Incorrect interventions
Menconi 2007[88]	No usable outcomes
Miranda-padua 2014[94]	Incorrect interventions
Müller 2001[95]	Inappropriate comparison
Nakamura 2007[96]	Incorrect interventions
Nedrebo 2002[98]	Incorrect interventions
Noh 2015[99]	Incorrect interventions
Orsini 2012[100]	Inappropriate comparison
Peixoto 2006[102]	Incorrect interventions
Peters 1995[103]	Incorrect interventions
Peters 1996[104]	No usable outcomes
Peters 1997[105]	Incorrect interventions
Pfeilschifter 1997[106]	Inappropriate comparison
Pirnat 2011[107]	Incorrect interventions
Pusuwan 2011[108]	Inappropriate comparison
Raber 2000[109]	Incorrect interventions
Reinwein 1993[110]	Inappropriate comparison
Rittmaster 1998[111]	Incorrect interventions
Rokni 2014[112]	SR checked for references
Romaldini 1983[113]	Incorrect interventions
Santos 2004[116]	NRS without adequate adjustment
Santos 2012[117]	Inappropriate comparison
Sapienza 2015[118]	Inappropriate comparison
Schneider 2005[119]	Inappropriate comparison
Singhal 2014[122]	Withdrawn Cochrane review
Taïeb 2016[123]	Incorrect interventions

Study	Exclusion reason
Thientunyakit 2010[125]	Inappropriate comparison
Tian 2001[126]	Not in English
Unalp 2009[129]	No usable outcomes
Walter 2006[131]	NRS without adequate adjustment
Wang 2016[132]	SR, checked for references
Weetman 1994[133]	Incorrect interventions
Witte 2000[134]	Incorrect interventions
Yousefi 2011[135]	Not in English
Yuan 2017[136]	SR, checked for references

J.1.2 Radioactive Iodine safety

Table 25: Studies excluded from the clinical review

Study	Exclusion reason
Angusti 2000[9]	No usable outcomes
Cevallos 1974[26]	Not minimum sample size
Chao 2009[27]	SR, references checked
Franklyn 1998[39]	No usable outcomes
Hall 1992[47]	No usable outcomes
Hall 1992[48]	Majority of radioactive iodine exposure not therapeutic
Hall 1993[51]	No usable outcomes
Hall 1995[49]	Outcomes reported elsewhere and included
Hall 1997[50]	Non-systematic review
Hieu 2012[55]	SR, references checked
Hoffman 1982[58]	No usable outcomes
Hoffman 1983[57]	Outcomes reported elsewhere and included
Holm 1980[61]	Outcomes reported elsewhere and included
Holm 2006[60]	Non-systematic review
Journy 2017[68]	Inappropriate population
Mctiernan 1984[87]	Inappropriate study design
Metso 2004[93]	No usable outcomes
Metso 2007[89]	Erratum, not relevant
Metso 2007[91]	Erratum, not relevant
Metso 2007[92]	No usable outcomes
Ron 1998[114]	No usable outcomes
Singer 2001[120]	Commentary only
Singer 2001[121]	No usable outcomes
Verburg 2011[130]	SR, references checked

J.2 Excluded health economic studies

J.2.1 Drugs vs Surgery vs Radioactive Iodine

Study	Exclusion reason

Study	Exclusion reason
Patel [101]	Not applicable, resource use and cost data from 2002

J.2.2 Radioactive Iodine safety

None

Appendix K: Research recommendations

K.1 Research question: Are there subgroups of people with Graves' thyrotoxicosis who have a particularly good response to antithyroid drugs?

Why this is important:

Antithyroid drugs (ATDs) are commonly used for treatment of Graves' disease. With a 12-18 months course of ATDs, there is about 50% chance for peoples with Graves' disease achieving a long-term remission. Previous studies have suggested patients with certain clinical characteristics are more likely to relapse following ATD. These characteristics, variably suggested by different studies, include male sex, young age, cigarette smoking, presence of large goitre, high levels of thyroid hormones at the time of diagnosis and high titres of TSH-receptor antibodies. However, most of these studies are small and retrospective in design, and these findings need confirmation by large prospective multi-centre studies. If the findings are confirmed, it will allow clinicians to stratify patients with Graves' hyperthyroidism who are unlikely to remain in remission following a course of ATD and offer early definitive treatments such as radioactive iodine or thyroidectomy.

Within the present guideline, the committee agreed that radioactive iodine should constitute the first line treatment option for adults with thyrotoxicosis/hyperthyroidism/Graves' disease according to both clinical and cost-effectiveness, but that for people in whom ATDs are particularly likely to achieve remission the, need for definitive treatment might be less. These were hypothesised by the committee to be people with milder, predominantly T3 thyrotoxicosis. However, no evidence was currently identified about any specific group of people who are likely to respond particularly well to ATDs. Further research is required to allow us to identify those people and allow clinicians to stratify patients with Graves' disease who are likely to remain in remission following the course of ATDs and avoid offering them a definitive treatment such as radioactive iodine or thyroidectomy.

Criteria for selecting high-priority research recommendations:

PICO question	Population: People with Graves' disease who are being treated with an antithyroid drug (ATD) Indicator: Absence of goitre, absence of thyroid eye disease, low titres of TSH receptor antibodies, low tires of free thyroid hormone levels at diagnosis, non-smoking mild thyrotoxicosis/ Graves' disease, T3 thyrotoxicosis… Comparator: Presence of goitre, presence of thyroid eye disease, high titres of TSH receptor antibodies, high titres of free thyroid hormone levels at diagnosis, smoking, non-mild Grave's disease… Outcome(s): hyperthyroidism relapse rate
Importance to patients or the population	This research will help to ascertain if simple clinical characteristics are useful in predicting the achievement of remission following a course of ATD. This will enable clinicians to stratify people with Graves' disease who are likely to achieve long-term remission after a course of ATD and those who are not and 'provide particular

	groups of people with the treatment they are most likely to benefit from' OR 'avoid definitive treatment for people who do not need it'.
Relevance to NICE guidance	This research will allow future guidelines to clearly recommend which people with Graves' disease should be offered ATDs as first line treatment instead of definitive treatment with radioactive iodine or thyroidectomy.
Relevance to the NHS	This research will provide clear evidence of the potential subgroup(s) of people with Graves' disease that could effectively be treated with an ATD. This will allow the identification of people who are likely to achieve long-term remission with a course of ATD, and avoid offering them definitive treatment in early course of the disease.
National priorities	Hyperthyroidism comes under the long-term condition directorate in the UK.
Current evidence base	Several retrospective single site studies have suggested various clinical characteristics, such as the presence of large goitre, high titres of free thyroid hormones at presentation, high titres of TSH receptor antibodies and smoking status are associated with the risk of relapse following a course of ATDs in patients with Graves' hyperthyroidism. However, no evidence about groups of people likely to respond particularly well to ATDs has been identified.
Equality	This recommendation is unlikely to impact on equality issues.
Study design	A multi-centre prospective observational study.
Feasibility	As Graves' disease is common, and ATDs are widely used in the UK for the treatment of Graves' disease, a multi-centre prospective observational study is feasible. A key challenge will be differences in clinical practice, in terms of regimes and duration of ATD, between different centres in the UK.
Other comments	
Importance	Medium: the research is relevant to the recommendations in the guideline, but the research recommendations are not key to future updates.

K2. Research question: What is the long-term clinical and cost effectiveness, including safety, of radioactive iodine for thyrotoxicosis?

Why this is important:

Radioactive iodine (I-131 NaI) is used to treat benign thyroid disease in approximately 10,000 patients in the UK each year by delivering absorbed doses (radiation) preferentially to the thyroid. This treatment is used globally with and is considered clinically effective. Despite the large number of patients treated with radioactive iodine over the past 50 years, there are still questions concerning the medium and longer-term effects and in particular the potential impact of exposure to low doses of radiation. A registry would enable the long-term effects of

radiation to be recorded and in time would provide definitive answers regarding the association of radiation and its specific dose with medium and long term risks. This would inform treatment protocols and would potentially provide reassurance to patients.

Criteria for selecting high-priority research recommendations:

PICO question	Population: Patients receiving radio-iodine for benign thyroid disease Intervention(s): Radio-iodine (RAI) therapy Comparison: The population not receiving RAI therapy Outcome(s): Neoplasia, fertility, quality of life, morbidity, death
Importance to patients or the population	The registry will be used to develop and refine our understanding of the risks and benefits associated with RAI therapy. This would help patients to make informed choices and place the risks / benefits of RAI into context.
Relevance to NICE guidance	Registry development would help to establish the role of RAI in the management of benign thyroid disease.
Relevance to the NHS	RAI therapy is relatively cost-effective when compared to interventions like surgery. There may be associated benefits from a financial and resource perspective.
National priorities	The NHS Five Year Forward View (2014) aims to address variations in treatment and outcomes. A register of treatments and outcomes would enable this information to be collected.
Current evidence base	RAI safety review did not support an association of RAI with increased risk of malignancy, however results have been largely based on older studies using lower RAI doses than those currently used in the UK. The general public and health professionals are not clear about the risks and benefits of RAI therapy. A registry documenting outcomes following RAI treatment in the medium and long term according to current practice will provide greater clarity in this area.
Equality	A registry might be of particular benefit in the context of young people receiving RAI who statistically will have more life-years ahead of them and an increased theoretical risk of health issues such as neoplasia as a result.
Study design	A central registry of all patients receiving RAI would be established. The key national bodies including those in the field of medical physics would agree to submit data on a regular basis. This data could be linked at national level to cancer registries / cause of death and patients will be asked to consent to being contacted about studies in areas such as QOL at a later stage. In the absence of consent, anonymised data will still be linked to long term morbidity / mortality data.
Feasibility	Collecting this data is a long-term project but is relatively inexpensive. A key issue would be to ensure high ascertainment.
Other comments	Radioactive iodine treatment for benign thyroid disease is performed widely. A registry could therefore have international impact.
Importance	• The research is important to quantify the risks associated with RAI therapy in greater detail.