National Institute for Health and Care Excellence

1

Draft

Thyroid Cancer

[I] Evidence review for thyrotropin alfa

NICE guideline

Evidence reviews underpinning recommendations 1.3.10 to 1.3.13 in the NICE guideline

June 2022

Draft for Consultation

These evidence reviews were developed by the National Guideline Centre

1

Disclaimer

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

Local commissioners and/or 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 All NICE guidance is subject to regular review and may be updated or withdrawn.

Copyright

ISBN:

Contents

1

1 Pretherapeutic thyrotropin alfa

1.1 Review question

1.1.1 What is the clinical and cost effectiveness of radioactive iodine with withdrawal of thyroid hormone replacement versus radioactive iodine with thyrotropin alfa?

1.1.2 Introduction

The uptake of radioactive iodine (RAI) is dependent on several factors but is primarily driven by thyroid stimulating hormone (TSH). Historically, total thyroid hormone withdrawal (THW) requiring a patient to stop thyroid hormone replacement for up to 4 weeks, has been the standard method of preparation for patients receiving RAI ablation, to allow the TSH to rise and therefore optimise RAI uptake. An alternative to stopping thyroid hormone replacement is the use of thyrotropin alfa, also known as recombinant Human TSH (rHTSH). Thyrotropin alfa is a synthetic form of thyroid stimulating hormone, which stimulates the thyroid tissue. This requires two intra-muscular injections on the two days before administration of the RAI. The easier administration and the avoidance of THW has resulted in this being accepted in clinical practice as preferred preparation for RAI treatment. This review investigates the evidence behind and the cost effectiveness of this approach.

1.1.3 Summary of the protocol

For full details see the review protocol in Appendix A.

Table 1: PICO characteristics of review question

Population	Inclusion: People aged 16 or over who have had thyroidectomy for differentiated thyroid cancer, and who are deemed suitable for RAI ablation/treatment. Exclusion: Children under 16
Intervention(s)	• radioactive iodine ablation/treatment with prior withdrawal of thyroid hormone replacement • radioactive iodine ablation/treatment with prior preparation with thyrotropin alfa
Comparison(s)	• Each other • radioactive iodine ablation/treatment with neither of the above two uptake-stimulating strategies
Outcomes	• mortality • quality of life (any validated scales) • local cancer progression (increase in size/number of tumours) • incidence of distant metastases • cancer recurrence • successful ablation • Second primary malignancy Longest available follow up in the studies.
Study design	• Systematic reviews

- RCTs

Non-randomised studies (any controlled designs, such as prospective/retrospective cohorts and case-control studies, with evidence of adjustment for biologically plausible confounders) will be included for one/both strata (ablation/treatment) if there are no RCTs in one/both strata.

1.1.4 Methods and process

This evidence review was developed using the methods and process described in Developing NICE guidelines: the manual. Methods specific to this review question are described in the review protocol in appendix A and the methods document.

Declarations of interest were recorded according to NICE's conflicts of interest policy.

1.1.5 Effectiveness evidence

1.1.5.1 Included studies

Eleven randomized controlled studies were included in the review;[10 12, 14-16, 21, 25, 33, 37 7, 39] which are summarised in Table 2 below. The studies compared radioiodine ablation with withdrawal of levothyroxine to radioiodine ablation with thyrotropin alfa. Evidence from these studies is summarised in the clinical evidence summary below (Table 2).

All papers were deemed to fit into the ablation stratum. There were several quality of life outcomes that showed heterogeneity, and that were therefore subject to exploratory sub-group analyses using the 3 sub-grouping strategies outlined in the protocol: use of dietary restrictions, TSH levels and RAI activity levels. The former two strategies were not useful as no studies reported dietary restrictions and all studies reported similar TSH levels in the THW group (>30 mU/L). However, for 3 quality of life outcomes, sub-grouping according to RAI activity levels helped to resolve heterogeneity. For these outcomes, therefore, the outcomes have been split according to studies where the activity was at 3.7 Gbq or studies where the activity was mixed (1.1/3.7 Gbq).

One Cochrane review (Ma 2010[23]) was excluded due to different protocol outcomes. The included studies were checked of which two were included in this review and the other two did not have any relevant outcomes.

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.

1.1.5.2 Excluded studies

See the excluded studies list in Appendix I.

1.1.6 Summary of studies included in the effectiveness evidence

Table 2: Summary of studies included in the evidence review

Study	Intervention and comparison	Population	Outcomes	Comments
Chianelli 2009[10]	**RAI + Withdrawal:** Radioactive iodine ablation - with prior withdrawal of thyroid hormone replacement. Twenty-one patients were treated with 131I in the hypothyroid state; L-T4 was stopped for 37 days; from the 3rd to 22nd day after L-T4 withdrawal patients were treated with T3. Patients received 131I (2.02±0.22 GBq; 54.6± 5.9 mCi) 42–180 days after surgery. L-T4 was then given again the day after administration of 131I (n=21) **RAI + rhTSH:** Radioactive iodine ablation - with prior preparation with thyrotropin alfa. Twenty-one patients were treated with 131I following the administration of rhTSH the therapeutic activity of 131I (1.97±0.18 GBq; 53.2±4.9 mCi) was administered 24 h after the last injection of rhTSH (0.9 mg i.m. for two consecutive days); L-T4 was never stopped during treatment. The time between thyroidectomy and 131I treatment was 42–180 days (n=21)	All patients had papillary cancer or minimally invasive follicular cancer, with a tumour node metastases stage pT1, larger than 1 cm or less than 1 cm if in the presence of multiple foci and could be considered patients at low risk of recurrence Age - Mean (SD): Withdrawal: 48±9.9; rhTSH: 46.1±12.3 RCT Italy	• Successful ablation	

Study	Intervention and comparison	Population	Outcomes	Comments
	(Follow up: 6 months)			
Emmanouilidis 2009[14]	**RAI + rhTSH:** Radioactive iodine ablation - with prior preparation with thyrotropin alfa. RhTSH participants received their first RAT on first hospitalization. rhTSH with a biological potency of 10 U/mg of protein was used according to the manufacturer's instructions. Each vial containing 0.9 mg of rhTSH-alfa was dissolved in 1.2 ml of water for injection and administered by the i.m. route to the gluteal region 48 and 24 h before RAT. After iodine uptake was confirmed by neck scan with 100 MBq 131I, the ablative activity of 3700 MBq 131I was administered orally (n=13) **RAI + Withdrawal:** Radioactive iodine ablation - with prior withdrawal of thyroid hormone replacement. patients in L-T4 abstinence group were discharged from the surgery ward and, while in a state of distinctive hypothyroidism, were re-hospitalized for the first RAT within 4–6 weeks after thyroidectomy. After iodine uptake was confirmed by neck scan with 100 MBq 131I, the ablative	Patients with a diagnosis of DTC or from patients that were thyroidectomized due to multinodular struma and who had a coincidental histology of DTC Age - Mean (SD): rhTSH: 45.2±16.5; Withdrawal: 54.8±12.8. RCT Germany	• Successful ablation • Cancer recurrence	

Study	Intervention and comparison	Population	Outcomes	Comments
	activity of 3700 MBq 131I was administered orally (n=12) (follow up: approximately 4 months after intervention)			
Emmanouilidis 2013[15]	**RAI + rhTSH**: Radioactive iodine ablation - with prior preparation with thyrotropin alfa. RhTSH patients received their first RAT on first hospitalization. RhTSH with a biological potency of 10 U/mg of protein was used according to the manufacturer's instructions 48 h and 24 h before RAT. After iodine uptake was confirmed by neck scan with 100Milli-Becquerel (MBq) 131I, the ablative activity of 3700MBq 131Iwas administered orally. (n=24) **RAI + Withdrawal:** Radioactive iodine ablation - with prior withdrawal of thyroid hormone replacement. patients in the L-T4 withdrawal group were discharged from hospital and readmitted for the first RAT within 4–6weeks after thyroidectomy while in a state of distinctive hypothyroidism. After iodine uptake was confirmed by neck scan with 100Milli-Becquerel (MBq) 131I, the ablative activity of	Patients with differentiated thyroid cancer awaiting radioiodine ablation therapy. Age - Median (range): rhTSH: 50 (17-66); Withdrawal: 58 (30-73). RCT Germany	• Cancer recurrence	

Study	Intervention and comparison	Population	Outcomes	Comments
	3700MBq 131Iwas administered orally (n=20)			
ESTIMABL1 trial: Schlumberger 2012[37] and Borget, 2015[7]	**RAI + rhTSH** Radioactive iodine ablation - with prior preparation with thyrotropin alfa. All patients underwent total thyroidectomy. 30 and 120 days after surgery, patients received levothyroxine therapy for at least 28 days (or levotri-iodothyronine therapy for 14 days). Recombinant human thyrotropin was administered during treatment with thyroid hormone, at a dose of 0.9 mg intra-muscularly on 2 consecutive days, and radioiodine was administered on the day after the second injection one of two 131I activities (1.1 GBq or 3.7 GBq). (n=374) **RAI + Withdrawal:** Radioactive iodine ablation - with prior withdrawal of thyroid hormone replacement. All patients underwent total thyroidectomy. Thyroid-hormone withdrawal consisted of discontinuation of levothyroxine treatment for at least 28 days (or levotriiodothyronine treatment withdrawal for 14 days), with administration of radioiodine when the serum thyrotropin concentration was higher than 30	Patients aged 18 years or older, low risk differentiated thyroid carcinoma (papillary or follicular, excluding aggressive histologic subtypes) Age - Mean (SD): rhTSH: 1.1GBq 51±13; 3.7GBq 48±14; Withdrawal: 1.1GBq 49±13; 3.7GBq 49±14. RCT France	• Successful ablation • Quality of life	

Study	Intervention and comparison	Population	Outcomes	Comments
	mIU per liter. radioiodine was administered at one of two 131I activities (1.1 GBq or 3.7 GBq). (n=378)			
HiLo Trial: Mallick 2012[25] merged with Dehbi 2019[12]	**RAI + rhTSH:** Radioactive iodine ablation - with prior preparation with thyrotropin alfa. Thyrotropin alfa was administered on each of the 2 days before ablation by intramuscular injection (0.9 mg) Radioactive iodine-131 was administered at a dose of 1.1 GBq (n=110) or 3.7 GBq (n=109), depending on the study group. (n=219) **RAI + Withdrawal:** Radioactive iodine ablation - with prior withdrawal of thyroid hormone replacement. Among the patients undergoing thyroid hormone withdrawal, thyroxine (average dose, 200 µg per day) was discontinued 4 weeks before ablation in 11 patients, and triiodothyronine (average dose, 60 µg per day) was discontinued for 2 weeks in 204 patients; Radioactive iodine-131 was administered at a dose of 1.1 GBq or 3.7 GBq, depending on the study group. (n=219) (follow up 3 – 9 months post intervention)	Patients aged 16 to 80 years, a performance status of 0 to 2, histological confirmation of differentiated thyroid cancer (including Hürthle-cell carcinoma) requiring radioiodine ablation; tumour stage T1 to T3 with the possibility of lymph-node involvement but no distant metastasis and no microscopical residual disease Age - Median (range): rhTSH: 44 (20-82) / 44 (21-76); Withdrawal: 45 (17-73) / 43 (18-77). RCT UK	• Successful ablation • Cancer recurrence • Quality of life	

Study	Intervention and comparison	Population	Outcomes	Comments
Lee 2010[21]	**RAI + rhTSH:** Radioactive iodine ablation - with prior preparation with thyrotropin alfa. All patients underwent total thyroidectomy with central compartment neck dissection. After the operation, all patients began treatment with TSH supressing dose of LT4 (levothyroxine 2µg / kg) after at least 30 days of LT4 supplementation. In the rhTSH group, each patient received two injections of rhTSH: 0.9mg IM at 24 hours and 48 hours before the administration of the RI therapeutic dose using low dose (30 mCi / 1.11GBq) radioiodine treatment. (n=69) **RAI + Withdrawal:** Radioactive iodine ablation - with prior withdrawal of thyroid hormone replacement. All patients underwent total thyroidectomy with central compartment neck dissection. After the operation, all patients began treatment with TSH supressing dose of LT4 (levothyroxine 2µg / kg) after at least 30 days of LT4 supplementation. Those in the T4 withdrawal group discontinued LT4 for 4 weeks. Remnant ablation using low dose (30 mCi / 1.11GBq) radioiodine treatment.	Patients with newly diagnosed disseminated thyroid cancer, more than 18 years old, who had recently undergone total or near total thyroidectomy with central compartment neck dissection. Age - Mean (SD): rhTSH: 46.7 ± 9.8; Withdrawal: 50.1 ± 6.8. RCT South Korea	• Successful ablation • Incidence of distant metastases	

Study	Intervention and comparison	Population	Outcomes	Comments
	(n=89)			
Pacini 2006[33] merged with Hanscheid 2006[16]	**RAI + rhTSH:** Radioactive iodine ablation - with prior preparation with thyrotropin alfa. Patients in the euthyroid group received l-thyroxine therapy for 4–6 wk until their serum TSH concentration was 5 mU/liter or less. Then 0.9 mg rhTSH was administered on 2 consecutive days; 24 h after rhTSH, 3.7 GBq (100 mCi) 131I was administered. (n=33) **RAI + Withdrawal:** Radioactive iodine ablation - with prior withdrawal of thyroid hormone replacement. Patients randomized to the hypothyroid group did not receive thyroid hormone therapy postoperatively. The serum TSH concentration was reassessed at 4–6 week until the patient's TSH was greater than 25mU/litre. The patients received a 3.7GBq(100 mCi) 131. (n=30)	Patients were 18 years or older with newly diagnosed differentiated papillary or follicular thyroid carcinoma, the sole previous treatment for which had been total or near-total thyroidectomy within 2 weeks before enrolment. Age - Mean (SD): Withdrawal: 43.2 (12.5); rhTSH: 44.5 (12.2).	• Successful ablation • Quality of life	
Taieb 2009[39]	**RAI + Withdrawal:** Radioactive iodine ablation - with prior withdrawal of thyroid hormone replacement. Patients were discharged from the department of endocrine surgery with levothyroxine supplementation (2µg/kg). One-week later patients	Aged ≥18 years, newly diagnosed well differentiated papillary or follicular carcinoma in patients who had total thyroidectomy (one stage or two stage) Age - Mean (SD): Withdrawal: 49 ± 11.8;	• Successful ablation • Incidence of distant metastases • Quality of life	

Study	Intervention and comparison	Population	Outcomes	Comments
	were randomized into the hypo group in which patients discontinued L-T4 for 5 weeks. All patients received 3.7GBq activity at 6 weeks post surgery. (n=37) **RAI + rhTSH:** Radioactive iodine ablation - with prior preparation with thyrotropin alfa. Patients were discharged from the department of endocrine surgery with levothyroxine supplementation (2µg/kg). One-week later patients were randomized into the rhTSH group in which patients continued to take L-T4 and received rhTSH (two 0.9mg IM injections on two consecutive days as ambulatory patients) 1 - 2 weeks later. Both injections were performed at the institution to ensure injection and TSH peak was validated. All patients received 3.7GBq activity at 2 - 3 weeks post-surgery. (n=37)	rhTSH: 45.5 ± 15.6. RCT France		

1

See Appendix D for full evidence tables.

1.1.7 Summary of the effectiveness evidence

Table 3: Clinical evidence summary: Radioiodine ablation with withdrawal of levothyroxine compared to radioiodine ablation with thyrotropin alfa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with comparator (RhTSH)	Risk difference between Withdrawal and RhTSH (95% CI)
Successful ablation (Tg<0.2ng/ml)	359 (1 study) 3 months	⊕⊕⊖⊖ LOW1 due to risk of bias	RR 0.98 (0.91 to 1.07)	Moderate	
				876 per 1000	18 fewer per 1000 (from 79 fewer to 61 more)
Successful ablation (Tg<0.2ng/ml) and <0.1 WBS%	421 (1 study) 3 months	⊕⊕⊕⊖ MODERATE1 due to risk of bias	RR 1 (0.92 to 1.07)	Moderate	
				871 per 1000	0 fewer per 1000 (from 70 fewer to 61 more)
Successful ablation (Tg<1ng/ml)	850 (3 studies) 6-9 months	⊕⊕⊕⊖ MODERATE1 due to risk of bias	RR 1 (0.97 to 1.04)	Moderate	
				941 per 1000	0 fewer per 1000 (from 28 fewer to 37 more)
Successful ablation (no visible uptake)	260 (3 studies) 6-12 months	⊕⊕⊕⊕ HIGH	RR 1.05 (0.97 to 1.14)	Moderate	
				905 per 1000	45 more per 1000 (from 27 fewer to 127 more)
Successful ablation (Tg<0.8µg/l + <0.1% WBS uptake)	71 (1 study) 9 months	⊕⊕⊕⊕ HIGH	RR 1.09 (0.96 to 1.24)	Moderate	
				889 per 1000	80 more per 1000 (from 36 fewer to 213 more)
Complete Ablation	684 (1 study) 6-10 months	⊕⊕⊖⊖ LOW1	RR 1.01	Moderate	
				917 per 1000	9 more per 1000 (from 28 fewer to 55 more)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with comparator (RhTSH)	Risk difference between Withdrawal and RhTSH (95% CI)
		due to risk of bias	(0.97 to 1.06)		
Visible uptake <0.1%	481 (2 studies) 6-9 months	⨁⨁⨁⨁ HIGH	RR 0.98 (0.93 to 1.04)	Moderate	
				594 per 1000	12 fewer per 1000 (from 42 fewer to 24 more)
Lymph node metastases	229 (2 studies) 9-12 months	⨁⨀⨀⨀ VERY LOW1,2 due to risk of bias, imprecision	RR 0.84 (0.2 to 3.52)	Moderate	
				22 per 1000	4 fewer per 1000 (from 18 fewer to 55 more)
Cancer recurrence	503 (3 studies) up to 4.5 years	⨁⨁⨀⨀ LOW2 due to imprecision	RR 0.72 (0.38 to 1.37)	Moderate	
				60 per 1000	17 fewer per 1000 (from 37 fewer to 22 more)
Thyroglobulin levels (ng/ml)	183 (2 studies) 12 months - 2.5 years	⨁⨁⨁⨁ HIGH2		The mean thyroglobulin levels (ng/ml) in the control groups was 0.12 ng/ml	The mean thyroglobulin levels (ng/ml) in the intervention groups was 0.04 higher (0.01 to 0.07 higher)
SF-36 score (mental component) Scale from: 0 to 100.	838 (3 studies) 1-4 months	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		The mean SF-36 score (mental component) in the control groups* was 44.6	The mean SF-36 score (mental component) in the intervention groups was 3.75 lower (6.13 lower to 1.38 lower)
SF-36 score (physical component) Scale from: 0 to 100.	838 (3 studies) 1-4 months	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		The mean SF-36 score (physical component) in the control groups* was 49.8	The mean SF-36 score (physical component) in the intervention groups was 5.36 lower (7.13 lower to 3.60 lower)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with comparator (RhTSH)	Risk difference between Withdrawal and RhTSH (95% CI)
SF-36 (physical functioning score) Scale from: 0 to 100.	838 (3 studies) 1-4 months	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision, inconsistency		The mean SF-36 (physical functioning score) in the control groups* was 85.3	The mean SF-36 (physical functioning score) in the intervention groups was 10.32 lower (20.48 lower to 0.17 lower)
SF-36 (role physical) Scale from: 0 to 100.	838 (3 studies) 1-4 months	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision, inconsistency		The mean SF-36 (role physical) in the control groups* was 66.7	The mean SF-36 (role physical) in the intervention groups was 14.14 lower (33.09 lower to 4.82 higher)
SF-36 (bodily pain) Scale from: 0 to 100. SUBGROUPED TO MIXED 1.1/3.7 Gbq	438 (1 study) 4 months	⨁⨁⨁⊖ MODERATE1 due to risk of bias		The mean SF-36 (bodily pain) in the control groups was 5.4 (pre-post difference value)	The mean SF-36 (bodily pain) in the intervention groups was 0.10 higher (7.40 lower to 7.60 higher)
SF-36 (bodily pain) Scale from: 0 to 100. SUBGROUPED TO 3.7 Gbq	400 (2 studies) 1 month	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		The mean SF-36 (bodily pain) in the control groups was 72.2	The mean SF-36 (bodily pain) in the intervention groups was 8.80 lower (13.65 lower to 3.95 lower)
SF-36 (vitality) Scale from: 0 to 100. SUBGROUPED TO MIXED 1.1/3.7 Gbq	438 (1 study) 4 months	⨁⨁⨁⊖ MODERATE1 due to risk of bias		The mean SF-36 (vitality) in the control groups was 4.5 (pre-post difference value)	The mean SF-36 (bodily pain) in the intervention groups was 0.4 lower (6.40 lower to 5.60 higher)
SF-36 (vitality) Scale from: 0 to 100. SUBGROUPED TO 3.7 Gbq	400 (2 studies) 1 month	⨁⨁⊖⊖ LOW1,2 due to risk of		The mean SF-36 (vitality) in the control groups was 55.3	The mean SF-36 (vitality) in the intervention groups was 14.68 lower (19.07 lower to 10.28 lower)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with comparator (RhTSH)	Risk difference between Withdrawal and RhTSH (95% CI)
		bias, imprecision			
SF-36 (general health) Scale from: 0 to 100.	838 (3 studies) 1-3 months	⨁⨁⨁⊖ MODERATE1 due to risk of bias		The mean SF-36 (general health) in the control groups* was 66.1	The mean SF-36 (general health) in the intervention groups was 1.83 lower (4.66 lower to 1.00 higher)
SF-36 (social functioning score) Scale from: 0 to 100. SUBGROUPED TO MIXED 1.1/3.7 Gbq	438 (1 study) 4 months	⨁⨁⨁⊖ MODERATE1 due to risk of bias		The mean SF-36 (social functioning score) in the control groups was 7.7 (pre-post difference value)	The mean SF-36 (social functioning score) in the intervention groups was 1.1 higher (6.10 lower to 8.30 higher)
SF-36 (social functioning score) Scale from: 0 to 100. SUBGROUPED TO 3.7 Gbq	400 (2 studies) 1 month	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		The mean SF-36 (social functioning score) in the control groups was 76.1	The mean SF-36 (social functioning score) in the intervention groups was 13.33 lower (18.17 lower to 8.49 lower)
SF-36 (role - emotional score) Scale from: 0 to 100.	838 (3 studies) 1-3 months	⨁⨁⊖⊖ LOW1,2 due to risk of bias, inconsistency		The mean SF-36 (role - emotional score) in the control groups* was 67.8	The mean SF-36 (role - emotional score) in the intervention groups was 8.13 lower (15.88 lower to 0.38 lower)
SF-36 (mental health score) Scale from: 0 to 100.	838 (3 studies) 1-3 months	⨁⨁⊖⊖ LOW1,2 due to risk of bias, inconsistency		The mean SF-36 (mental health score) in the control groups was 68.5	The mean SF-36 (mental health score) in the intervention groups was 3.84 lower (9.06 lower to 1.39 higher)
EQ5D Utility score: Scale from 0-1	684 (1 study) 8 months	⨁⨁⨁⊖ MODERATE1 due to risk of bias		The mean EQ5D utility in the control groups was 0.849	The mean EQ5D utility score in the intervention groups was 0.02 lower (0.04 lower to 0.01 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with comparator (RhTSH)	Risk difference between Withdrawal and RhTSH (95% CI)
Physical Well-being Scale from: 0 to 28.	71 (1 study) ablation period	⨁⨁⨁⨀ MODERATE1 due to risk of bias		The mean physical well-being in the control groups was -0.62	The mean physical well-being in the intervention groups was 5.16 lower (7.24 to 3.08 lower)
Physical Well-being Scale from: 0 to 28.	71 (1 study) 3 months post ablation	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		The mean physical well-being in the control groups was 0.37	The mean physical well-being in the intervention groups was 1.95 lower (4.44 lower to 0.54 higher)
Physical Well-being Scale from: 0 to 28.	72 (1 study) 6 months post ablation	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		The mean physical well-being in the control groups was 0.14	The mean physical well-being in the intervention groups was 0.23 lower (2.32 lower to 1.86 higher)
Physical Well-being Scale from: 0 to 28.	71 (1 study) 9 months post ablation	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		The mean physical well-being in the control groups was -1.11	The mean physical well-being in the intervention groups was 0.42 higher (2.08 lower to 2.92 higher)
Social / Familial Well-being Scale from: 0 to 28.	71 (1 study) ablation period	⨁⨁⨁⨀ MODERATE1 due to risk of bias		The mean social / familial well-being in the control groups was -0.11	The mean social / familial well-being in the intervention groups was 4.89 lower (6.38 to 3.4 lower)
Social / Familial Well-being Scale from: 0 to 28.	71 (1 study) 3 months post ablation period	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		The mean social / familial well-being in the control groups was -0.32	The mean social / familial well-being in the intervention groups was 0.06 higher (1.54 lower to 1.66 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with comparator (RhTSH)	Risk difference between Withdrawal and RhTSH (95% CI)
Social / Familial Well-being Scale from: 0 to 28.	72 (1 study) 6 months post ablation	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		The mean social / familial well-being in the control groups was -0.15	The mean social / familial well-being in the intervention groups was 0.59 lower (2.88 lower to 1.7 higher)
Social / Familial Well-being Scale from: 0 to 28.	71 (1 study) 9 months post ablation	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		The mean social / familial well-being in the control groups was -0.45	The mean social / familial well-being in the intervention groups was 0.61 higher (1.12 lower to 2.34 higher)
Emotional Well-being Scale from: 0 to 24.	71 (1 study) ablation period	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		The mean emotional well-being in the control groups was 0.86	The mean emotional well-being in the intervention groups was 1.21 lower (2.75 lower to 0.33 higher)
Emotional Well-being Scale from: 0 to 24.	71 (1 study) 3 months post ablation	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		The mean emotional well-being in the control groups was 1	The mean emotional well-being in the intervention groups was 0.64 higher (1.11 lower to 2.39 higher)
Emotional Well-being Scale from: 0 to 24.	72 (1 study) 6 months post ablation	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		The mean emotional well-being in the control groups was 0.47	The mean emotional well-being in the intervention groups was 0.47 higher (1.42 lower to 2.36 higher)
Emotional Well-being Scale from: 0 to 24.	71 (1 study) 9 months post ablation	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		The mean emotional well-being in the control groups was 0.28	The mean emotional well-being in the intervention groups was 0.94 higher (0.92 lower to 2.8 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with comparator (RhTSH)	Risk difference between Withdrawal and RhTSH (95% CI)
Functional Well-being Scale from: 0 to 28.	71 (1 study) ablation period	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		The mean functional well-being in the control groups was -1	The mean functional well-being in the intervention groups was 1.49 lower (3.78 lower to 0.8 higher)
Functional Well-being Scale from: 0 to 28.	71 (1 study) 3 months post ablation	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		The mean functional well-being in the control groups was 0.89	The mean functional well-being in the intervention groups was 0.88 higher (1.59 lower to 3.35 higher)
Functional Well-being Scale from: 0 to 28.	72 (1 study) 6 months post ablation	⨁⨀⨀⨀ VERY LOW1,2 due to risk of bias, imprecision		The mean functional well-being in the control groups was 1.53	The mean functional well-being in the intervention groups was 0.59 higher (2 lower to 3.18 higher)
Functional Well-being Scale from: 0 to 28.	71 (1 study) 9 months post ablation	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		The mean functional well-being in the control groups was 0.83	The mean functional well-being in the intervention groups was 1.36 higher (0.98 lower to 3.7 higher)
Fatigue Scale from: 0 to 52.	71 (1 study) ablation period	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		The mean fatigue in the control groups was 0.86	The mean fatigue in the intervention groups was 1.21 lower (2.75 lower to 0.33 higher)
Fatigue Scale from: 0 to 52.	71 (1 study) 3 months post ablation	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		The mean fatigue in the control groups was 1	The mean fatigue in the intervention groups was 0.64 higher (1.11 lower to 2.39 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with comparator (RhTSH)	Risk difference between Withdrawal and RhTSH (95% CI)
Fatigue Scale from: 0 to 52.	72 (1 study) 6 months post ablation	⨁⨀⨀⨀ VERY LOW1,2 due to risk of bias, imprecision		The mean fatigue in the control groups was 0.47	The mean fatigue in the intervention groups was 0.47 higher (1.42 lower to 2.36 higher)
Fatigue Scale from: 0 to 52.	71 (1 study) 9 months post ablation	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		The mean fatigue in the control groups was 0.28	The mean fatigue in the intervention groups was 0.94 higher (0.92 lower to 2.8 higher)
Facit-F (TOI) Scale from: 0 to 52.	71 (1 study) ablation period	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		The mean facit-f (toi) in the control groups was -2.59	The mean facit-f (toi) in the intervention groups was 12.47 lower (20.05 to 4.89 lower)
Facit-F (TOI) Scale from: 0 to 108.	71 (1 study) 3 months post ablation	⨁⨀⨀⨀ VERY LOW1,2 due to risk of bias, imprecision		The mean facit-f (toi) in the control groups was 2.4	The mean facit-f (toi) in the intervention groups was 0.67 higher (8.67 lower to 10.01 higher)
Facit-F (TOI) Scale from: 0 to 108.	72 (1 study) 6 months post ablation	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		The mean facit-f (toi) in the control groups was 2.42	The mean facit-f (toi) in the intervention groups was 2.76 higher (6.21 lower to 11.73 higher)
Facit-F (TOI) Scale from: 0 to 108.	71 (1 study) 9 months post ablation	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		The mean facit-f (toi) in the control groups was -0.51	The mean facit-f (toi) in the intervention groups was 5.81 higher (3.48 lower to 15.1 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with comparator (RhTSH)	Risk difference between Withdrawal and RhTSH (95% CI)
FACT-G (total score) Scale from: 0 to 108.	71 (1 study) ablation period	⨁⨁⨁⨀ MODERATE1 due to risk of bias		The mean fact-g (total score) in the control groups was 1.63	The mean fact-g (total score) in the intervention groups was 11.45 lower (17.58 to 5.32 lower)
FACT-G (total score) Scale from: 0 to 108.	71 (1 study) 3 months post ablation	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		The mean fact-g (total score) in the control groups was 2.37	The mean fact-g (total score) in the intervention groups was 0.46 higher (5.43 lower to 6.35 higher)
FACT-G (total score) Scale from: 0 to 108.	72 (1 study) 6 months post ablation	⨁⨀⨀⨀ VERY LOW1,2 due to risk of bias, imprecision		The mean fact-g (total score) in the control groups was 1.85	The mean fact-g (total score) in the intervention groups was 0.03 lower (6.73 lower to 6.67 higher)
FACT-G (total score) Scale from: 0 to 108.	71 (1 study) 9 months post ablation	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		The mean fact-g (total score) in the control groups was -0.1	The mean fact-g (total score) in the intervention groups was 4.9 higher (0.71 lower to 10.51 higher)
Facit-F (total score) Scale from: 0 to 160.	71 (1 study) ablation period	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		The mean facit-f (total score) in the control groups was -4.05	The mean facit-f (total score) in the intervention groups was 12.21 lower (22.25 to 2.17 lower)
Facit-F (total score) Scale from: 0 to 160.	71 (1 study) 3 months post ablation period	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		The mean facit-f (total score) in the control groups was 4.26	The mean facit-f (total score) in the intervention groups was 3.51 higher (6.54 lower to 13.56 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with comparator (RhTSH)	Risk difference between Withdrawal and RhTSH (95% CI)
Facit-F (total score) Scale from: 0 to 160.	72 (1 study) 6 months post ablation	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		The mean facit-f (total score) in the control groups was 1.4	The mean facit-f (total score) in the intervention groups was 3.88 higher (6.58 lower to 14.34 higher)
Facit-F (total score) Scale from: 0 to 160.	71 (1 study) 9 months post ablation	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		The mean facit-f (total score) in the control groups was 0.8	The mean facit-f (total score) in the intervention groups was 10.33 higher (0.28 to 20.38 higher)

1 Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

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

3 Downgraded by 1 increment if the I^2 value was between 50% and 75% and downgraded by 2 increments if the I^2 value was over 75%.

*This mean value only included the two studies with the post test values and does not include the study with the post-pre values (which would otherwise skew the mean).

**The MIDs for binary outcomes were based on default OR, RR or HR values of 0.8 or 1.25. For continuous variables, the MIDs were based on the default value of ±0.5 x the median standard deviation (sd) in the control group. The median control group sd, together with the MID for all continuous variables, have been tabulated below:

Outcome	Control group median sd	MID
Thyroglobulin levels (ng/ml)	0.16	0.08
SF-36 score (mental component)	12	6
SF-36 score (physical component)	8	4
SF-36 (physical functioning score)	18.3	9.15
SF-36 (role physical)	38.9	19.45
SF-36 (bodily pain) SUBGROUPED TO MIXED 1.1/3.7 Gbq	40.04	20.02
SF-36 (bodily pain) SUBGROUPED TO 3.7 Gbq	23.3	11.65
SF-36 (vitality) SUBGROUPED TO MIXED 1.1/3.7 Gbq	32.02	16.01
SF-36 (vitality) SUBGROUPED TO 3.7 Gbq	22.25	11.12
SF-36 (general health)	20.8	10.4
SF-36 (social functioning score) SUBGROUPED TO MIXED 1.1/3.7 Gbq	38.44	19.22

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with comparator (RhTSH)	Risk difference between Withdrawal and RhTSH (95% CI)
SF-36 (social functioning score) SUBGROUPED TO 3.7 Gbq			22.2	11.1	
SF-36 (role - emotional score)			40.04	20.02	
SF-36 (mental health score)			21	10.5	
EQ5D Utility score:			0.173	0.0865	
Physical Well-being 0			2.71	1.35	
Physical Well-being 3m			4.4	2.2	
Physical Well-being 6m			3.94	1.97	
Physical Well-being 9m			4.86	2.43	
Social / Familial Well-being 0			1.7	0.85	
Social / Familial Well-being 3m			3.18	1.59	
Social / Familial Well-being 6m			3.2	1.6	
Social / Familial Well-being 9m			3.24	1.62	
Emotional Well-being 0			2.39	1.2	
Emotional Well-being 3m			2.94	1.47	
Emotional Well-being 6m			2.14	1.07	
Emotional Well-being 9m			3.1	1.55	
Functional Well-being 0			3.66	1.83	
Functional Well-being 3m			3.76	1.88	
Functional Well-being 6m			3.45	1.73	
Functional Well-being 9m			4.67	2.3	
Fatigue 0			2.39	1.2	
Fatigue 3m			2.94	1.5	
Fatigue 6m			2.14	1.07	
Fatigue 9m			3.1	1.55	
Facit-F (TOI) 0			12.89	6.45	
Facit-F (TOI) 3m			16.4	8.2	
Facit-F (TOI) 6m			16.26	8.13	
Facit-F (TOI) 9m			18.6	9.3	
FACT-G (total score) 0			7.72	3.86	
FACT-G (total score) 3m			9.83	4.93	
FACT-G (total score) 6m			8.22	4.11	
FACT-G (total score) 9m			10.82	5.41	

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with comparator (RhTSH)	Risk difference between Withdrawal and RhTSH (95% CI)
Facit-F (total score) 0			15.83	7.92	
Facit-F (total score) 3m			18.91	9.45	
Facit-F (total score) 6m			18.24	9.12	
Facit-F (total score) 9m			20.35	10.17	

See Appendix F for full GRADE tables.

1.1.8 Economic evidence

1.1.8.1 Included studies

Four health economic studies with the relevant comparison were included in this review. [5 7 28 38 43] These are summarised in the health economic evidence profile below (**Table 4**) and the health economic evidence tables in Appendix H.

1.1.8.2 Excluded studies

Two economic studies relating to this review question were identified but were excluded due to limited applicability [5, 46] and the availability of more applicable evidence. [27] These are listed in Appendix I, with reasons for exclusion given.

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

1.1.9 Summary of included economic evidence

Table 4: Health economic evidence profile: Radioactive iodine with and without thyroid-stimulating hormone

Study	Applicability	Limitations	Other comments	Incremental cost	Incremental effects	Cost effectiveness	Uncertainty
Borget 2015 [7] ([France])	Partially applicable [(a)]	Minor limitations [(b)]	Within-RCT cost-utility analysis (ESTIMBAL trial/ Schlumberger 2012[37]) Cost-utility analysis (QALYs) Population: Adults who underwent total thyroidectomy for low risk differentiated thyroid cancer prior to radioiodine ablation Comparators: Endogenous stimulation of TSH with THW Exogenous stimulation of TSH with rhTSH Follow-up: 8 months	£582[(c)]	0.012 QALYs	£48,500 per QALY gained	Probability that Intervention 2 was cost effective (£20K/30K threshold): 1.5%/22% Uncertainty: When the cost of rhTSH was reduced by 30%, the probability that rhTSH was cost effective at a threshold of £ 42,830 was 70%.
Mernagh 2010 [28] ([Canada])	Partially applicable [(d)]	Potentially serious limitations [(e)]	Markov model adapted from Mernagh 2006 [27] Cost-utility analysis (QALYs) Population: Adults who underwent total thyroidectomy for low risk differentiated thyroid cancer prior to radioiodine ablation Comparators:	£51[(f)]	0.0576 QALYs	£890	Probability rhTSH cost effective (£20k/30k threshold): NR Uncertainty: Several sensitivity analyses were conducted. However, they included societal costs and therefore it was not possible to interpret these findings from the

Study	Applicability	Limitations	Other comments	Incremental cost	Incremental effects	Cost effectiveness	Uncertainty
			Endogenous stimulation of TSH with THW Exogenous stimulation of TSH with rhTSH Time horizon: 17 weeks				perspective of the healthcare system..
Sohn 2015[38] ([South Korea])	Partially applicable [g]	Potentially serious limitations [h]	Markov model based on Mernagh 2010 [28] Cost-utility analysis (QALYs) Population: Adults who underwent total thyroidectomy for low risk differentiated thyroid cancer prior to radioiodine ablation Comparators: Endogenous stimulation of TSH with THW Exogenous stimulation of TSH with rhTSH Time horizon: 17 weeks	£769[i]	0.036 QALYs	£21,357 per QALY gained	Probability rhTSH cost effective (£20k/30k threshold): NR Uncertainty: Inclusion of indirect costs (i.e. loss of productivity) resulted in an incremental cost of £18,848 per QALY gained.
Vallejo 2017[43] ([Spain])	Partially applicable [j]	Potentially serious limitations [k]	Markov model based on Mernagh 2010 [28] Cost-utility analysis (QALYs) Population: Adults who underwent total thyroidectomy for low risk differentiated thyroid cancer prior to radioiodine ablation Comparators:	-£640 [l]	0.048 QALYs	Dominant (greater QALY gain at a lower cost)	Probability rhTSH cost effective (£20k/30k threshold): NR Uncertainty: Assuming no difference between treatment arms in hospital length of stay resulted in an incremental cost of £1,057 per QALY gained.

Study	Applicability	Limitations	Other comments	Incremental cost	Incremental effects	Cost effectiveness	Uncertainty
			Endogenous stimulation of TSH with THW Exogenous stimulation of TSH with rhTSH Time horizon: 17 weeks				

Abbreviations: ICER = incremental cost-effectiveness ratio; RCT = randomized controlled trial; QALYs= quality-adjusted life years; rhTSH = recombinant human thyroid stimulating hormone; TSH = thyroid stimulating hormone; THW = thyroid hormone withdrawal.

(a) Comparators included four strategies, each combining one of two TSH stimulation methods and one of two radioactive iodine doses. Results were reported for each of the four trial arms and as averages across endogenous and exogenous intervention arms. French healthcare context. Utility values used to calculate QALYs were derived from EQ-5D scores using French tariff.

(b) Incremental QALY gain reported (0.013) differs from that calculated from reported total mean values for each intervention (0.012). Limited sensitivity analyses were conducted. Disclosures provided by authors were not identified online.

(c) 2013 French euros converted to 2013 UK pounds.[32]. Cost components incorporated: Intervention cost, fixed hospital costs (staff, equipment, overhead), variable hospital costs (resources required for radioiodine administration, rhTSH, radioiodine activity).

(d) Canadian healthcare context. Disaggregated direct and societal results reported for the base case but not sensitivity analyses. Utility weights estimated using SF-6D mapping algorithm.

(e) No intervention effect was applied based on results of equivalence study by Pacini 2006. Ontario was used as the reference province for resource use and unit costs. Incremental quality of life estimated from a single trial that was an outlier in the meta-analysis. Several assumptions were needed to model quality of life over time as only two data points were available.

(f) 2007 Canadian dollars converted to 2007 UK pounds.[32].Cost components incorporated: Intervention cost (2 ampoules of Thyrogen®), ablative dose of ^{131}I radioiodine, whole body scan using radioiodine, inpatient hospitalization days for patients receiving radioiodine ablation, initial and follow-up specialist visits (radiation oncologist), initial and follow-up general practitioner visits, laboratory tests (serum thyroglobulin count, thyroglobulin antibody test), daily T4 medication.

(g) Korean healthcare context. Utility weights estimated using SF-6D mapping algorithm. Incremental quality of life estimated from a single trial that was an outlier in the meta-analysis. Several assumptions were needed to model quality of life over time as only two data points were available.

(h) No intervention effect was applied based on results of equivalence study by Pacini 2006. Cost year not reported and assumed to be 2013 based on unit cost reference dates. Conflict of interest declaration was unclear - the supervising author is a medical advisor in Genzyme Corporation which funded the study. Incremental quality of life estimated from a single trial that was an outlier in the meta-analysis. Several assumptions were needed to model quality of life over time as only two data points were available.

(i) 2013 South Korean won converted to UK pounds.[32]. Cost components incorporated: Intervention cost (2-vial kit of Thyrogen), ablative dose of radioiodine, whole body scan using radioiodine, inpatient hospitalization days for patients receiving radioiodine ablation, specialist visit (radiation oncologist), practice nurse visit, laboratory tests (TSH quantification test, serum thyroglobulin count, thyroglobulin antibody test), weekly T4 and T3 medication.

(j) Spanish healthcare context. Utility weights estimated using SF-6D mapping algorithm.

(k) No intervention effect was applied based on results of equivalence study by Pacini 2006. Incremental quality of life estimated from a single trial that was an outlier in the meta-analysis. Several assumptions were needed to model quality of life over time as only two data points were available.

(l) 2015 Spanish euros converted to 2015 UK pounds.[32]. Cost components incorporated: Intervention cost (2-vial kit of Thyrogen), ablative dose of radioiodine, whole body scan using radioiodine, inpatient hospitalization days for patients receiving radioiodine ablation, specialist visit (radiation oncologist), practice nurse visit, laboratory tests (TSH quantification test, serum thyroglobulin count, thyroglobulin antibody test), weekly T4 and T3 medication

1.1.10 Economic model

A quality-of-life simulation model was developed to asses the cost-effectiveness of rhTSH compared to thyroid hormone withdrawal (THW) in England in people who received total thyroidectomy and are preparing for RAI. The full economic report can be viewed in the economic report published alongside the guideline.

Population and strategies

The population of the analysis was people in preparation to receive RAI and the two strategies compared were:

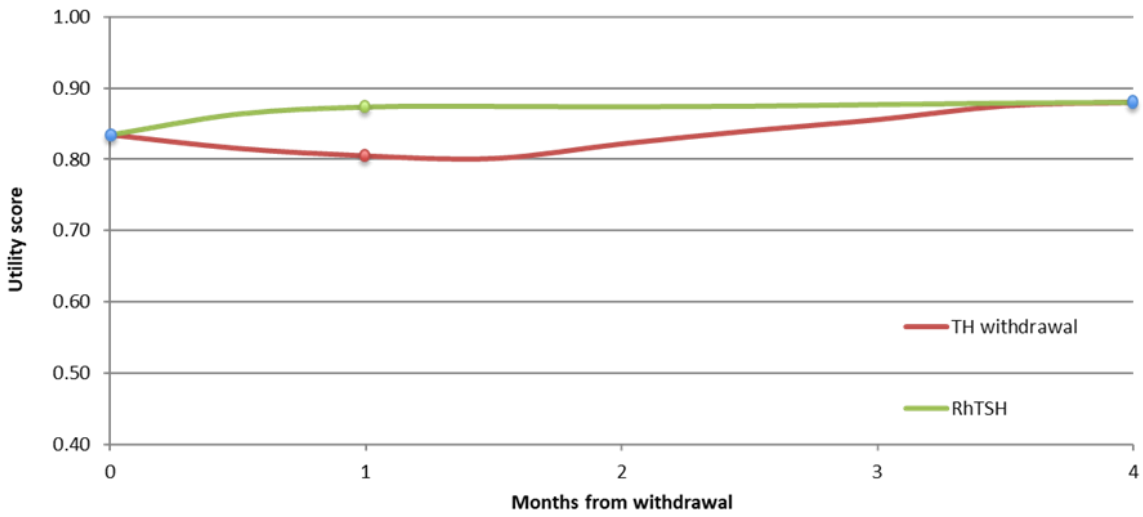
1. Exogenous TSH stimulation with recombinant human TSH using thyrotropin alpha
2. Endogenous TSH stimulation with thyroid hormone withdrawal (THW)

A cost-utility analysis was undertaken where quality-adjusted life years (QALYs) and costs from a current UK NHS and personal social services perspective were considered.

Model structure and data sources

- A quality-of-life simulation model was developed to estimate changes in quality of life among people receiving either THW or rhTSH
- The time horizon was set at 4 months and half as beyond this point in time no difference in quality of life or healthcare was observed. A cycle of a half-month was utilized to allow quality of life to vary in the two groups.
- Effectiveness data were estimated using a meta-analysis of all three clinical trials available[25, 33, 37]. SF-36 dimension scores were mapped into EQ-5D utility scores using Ara and Brazier algorithm[1]
- The utility curve estimated from Borget 2015[7] and based on ESTIMABL[37] was refitted using a meta-analysis of the three trials available. With this approach, EQ-5D utility scores follow the same distribution observed in the trial (see Figure 1)
- A proportion of people in the THW arm were assumed to switch to T3 before beginning withdrawal in one of the two main scenarios. In another scenario, everyone was assumed to receive T4. This is because the price of T3 in England is unusually high and expected to play a major role in the analysis
- People in the THW group are assumed to need additional healthcare services during the 4 weeks they experience withdrawal-induced hypothyroidism. These additional costs were estimated using the results of a survey on healthcare utilization during withdrawal-induced hypothyroidism[22]
- A threshold analysis on the level of adherence in THW group was conducted. Non-adherence was defined as the probability of someone showing up for RAI with a TSH level insufficient to receive the treatment. In this case, it was assumed they would receive rhTSH with Thyrotropin Alpha (TA).

Figure 1: Utility curves in rhTSH and THW groups using meta-analysed EQ-5d utility scores



Costs

- Pharmaceutical costs for T3, T4 and Thyrotropin Alpha (TA) were estimated using BNF[6] and prescription cost analysis database[17]
- Healthcare costs of RAI, endocrinology attendance and outpatient attendance were collected from the NHS Refence Costs 2019/2020[30]
- Cost of a GP visit was estimated using PSSRU[11]

Results

The two main scenarios were developed fully probabilistic.

- Scenario 1: around half of people in the THW group switch to T3 before beginning withdrawal using a proportion calculated from ESTIMBAL[37]
- Scenario 2: no one is assumed to switch to T3 and people assume only T4 throughout the duration of the analysis

The probabilistic results of the two main scenarios are presented in Table 5, Table 6 and Table 7.

Table 5: Probabilistic costs and QALYs in scenario 1

	THW	rhTSH	Difference (rhTSH – TWH)
Cost[(a)]	£1,191 (£1,162 to £1,224)	£1,515 (£1,506 to £1,526)	£323 (£292 to £351)
QALYs[(a)]	0.31 (0.27 to 0.36)	0.33 (0.27 to 0.38)	0.011 (0.003 to 0.021)

(a) Costs and QALYs are calculated per person

Table 6: Probabilistic costs and QALYs in scenario 2

	THW	rhTSH	Difference (rhTSH – TWH)
Cost[(a)]	£1,133 (£1,103 to £1,165)	£1,515 (£1,506 to £1,526)	£382 (£351 to £410)
QALYs[(a)]	0.31 (0.27 to 0.36)	0.33 (0.27 to 0.38)	0.012 (0.003 to 0.021)

(a) Costs and QALYs are calculated per person

Table 7: Probabilistic cost-effectiveness results

rhTSH vs THW	Scenario 1	Scenario 2
Cost per QALY	£27,315	£32,330
Probability rhTSH cost effective at £20,000 threshold	18%	7%
Probability rhTSH cost effective at £30,000 threshold	59%	43%

In both scenarios cost per QALY was above £20,000 although in Scenario 1 cost per QALY is below NICE threshold of £30,000. The probability of rhTSH being cost-effective at £20,000 threshold is 18% in Scenario 1 and 7% in Scenario 2.

Table 8 illustrates the results of the deterministic sensitivity analysis. RhTSH was found to be cost effective when a larger use of T3 was assumed and when QALYs estimation were based on Pacini 2006 trial[33] that, among the three trials, found the largest difference in quality of life between the two interventions. When historical prices from 2007 were assumed for T3, rhTSH was not cost-effective at a £30,000 threshold anymore.

Table 8: Deterministic scenario analyses results

	Incremental cost	Incremental QALYs	Cost per QALY
Scenario 1 (probabilistic)	£323	0.012	£27,315
Scenario 2 (probabilistic)	£382	0.012	£32,330
Give T3 to people for 2 weeks after withdrawal	£164	0.012	£13,914
Equal weight to each trial	£323	0.014	£22,769
Utilities based on Pacini 2006	£323	0.023	£13,776
Utilities based on ESTIMABL	£323	0.012	£27,562
Utilities based on HiLo	£323	0.009	£35,570
SF-6D utility score (ESTIMABL only)	£323	0.007	£48,777
2007 price for T3	£378	0.012	£32,021

Figure 2 and Figure 3 illustrate the results of the threshold analysis in scenarios 1 and 2. In scenario 1, rhTSH became cost-effective at £20,000 when between 1 and 2 out of 10 people do not have a sufficient level of TSH and need nan injection of Thyrotropin Alpha. In scenario 2, rhTSH becomes cost-effective only when adherence falls below 75%.

Figure 2: Threshold analysis on adherence in THW group (scenario 1)

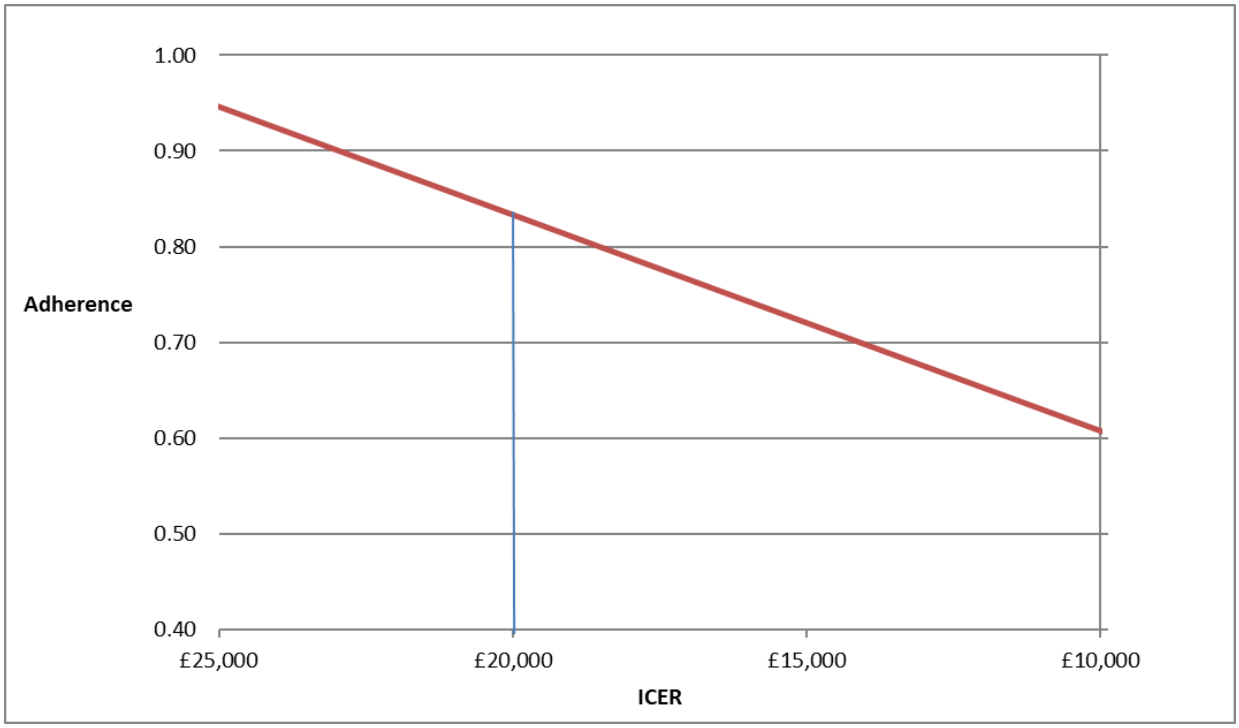
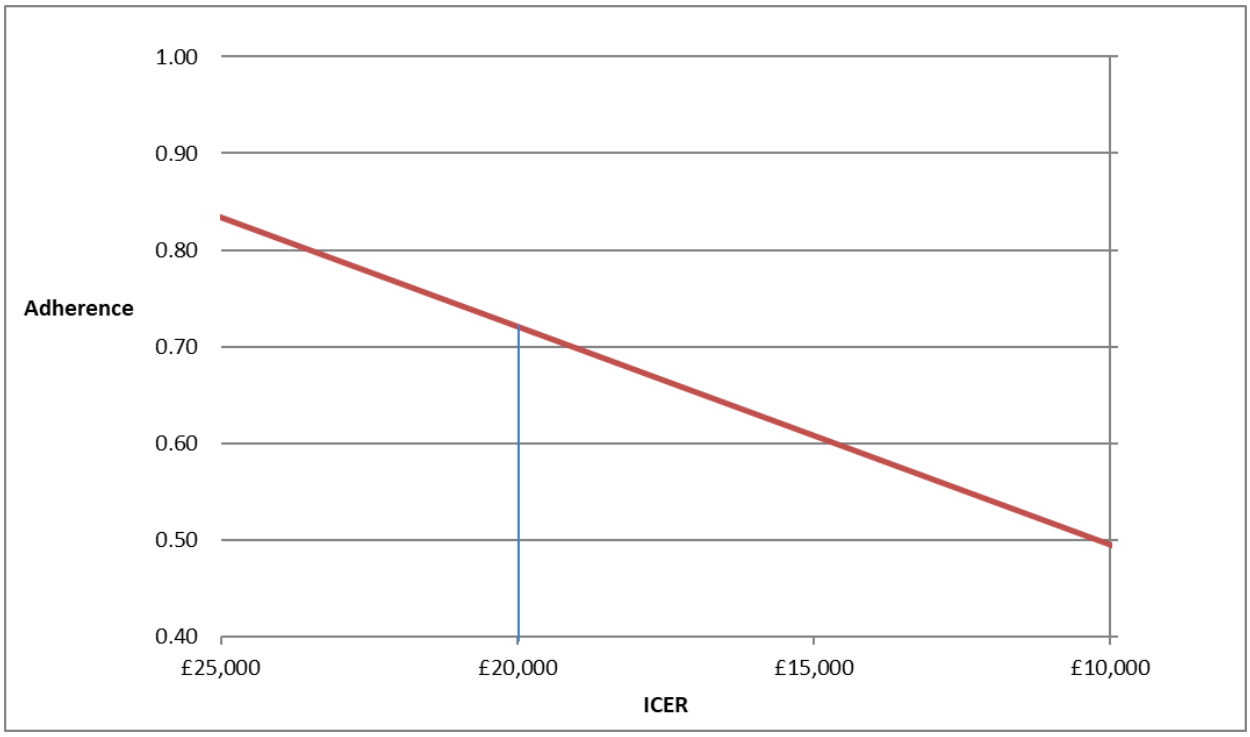


Figure 3: Threshold analysis on adherence in TWH group (scenario 2)



1.1.11 Economic evidence statements

- One cost-utility analysis found rhTSH not cost effective compared to THW. The analysis was assessed as partially applicable with minor limitations.
- Three cost-utility analysis found rhTSH cost-effective or dominant compared to THW. The analysis were assessed as partially applicable with potentially serious limitations.
- One original cost-utility analysis found rhTSH potentially cost-effective compared to THW in England (ICER: £23,002). The analysis was assessed as partially applicable with minor limitations

1.1.12 The committee's discussion and interpretation of the evidence

1.1.12.1 The outcomes that matter most

Critical outcomes selected by the committee for decision making were mortality, quality of life (any validated scales), local cancer progression (increase in size/number of tumours), incidence of distant metastases, cancer recurrence, successful ablation and second primary malignancy.

1.1.12.2 The quality of the evidence

Quality of evidence varied by outcome. The majority of outcomes were graded as low, with some moderate and very low and only a few graded as high. Most of the downgrading resulting from risk of bias and imprecision. Reasons for high risk of bias included lack of blinding and incomplete outcome data. Importantly, many of the outcomes used to make recommendations were those that were graded as moderate or high, and the committee agreed that the confidence in the evidence findings was strong.

1.1.12.3 Benefits and harms

Radioactive iodine ablation *with thyroid hormone withdrawal (THW)* and *RAI with thyrotropin alfa (rhTSH)* did not differ in longer term oncological outcomes such as successful ablation, lymph node metastases, cancer recurrence or thyroglobulin levels. However, clear benefits for rhTSH over THW were evident for well-being, social function, emotional function, general function and fatigue at the time of ablation, although these benefits were not sustained over time.

These short-term benefits for patients receiving rhTSH were not a surprise to the committee, who explained these effects through two mechanisms. Firstly, the use of rhTSH instead of THW will avoid hypothyroidism, thus side-stepping the deleterious effects of hypothyroidism on function and general quality-of-life in the peri-ablation period. Secondly, the avoidance of hypothyroidism will reduce impairment of renal function, which will facilitate more rapid excretion of radioactive iodine than otherwise. The more rapid excretion of radioactive iodine will reduce the total dose of absorbed radiation in patients prepared with rhTSH versus THW (even though the *administered* dose will be the same in both treatments) and should improve general health and well-being in the short term, as well as a quicker return to normal life. The committee questioned whether this reduction in absorbed radiation dose in those receiving rhTSH might confer reduced effectiveness, but this was regarded as unlikely, based on the evidence that successful ablation did not differ between rhTSH and THW.

Although the evidence review did not capture longer term outcomes, the committee considered that reducing absorbed doses of radioactive iodine may lead to a decrease in the risk of second malignancies. Lifetime prevalence of second malignancies after radioactive ablation was cited as 1 in 200, and it was presumed very plausible that reducing absorbed doses through using rhTSH may have long term benefits in terms of reducing malignancy risk.

The committee discussed the benefits and harms associated with both treatments. rhTSH enables people to return to normal activities within 2 or 3 days of treatment, whereas with THW is taken for 4 to 6 weeks before treatment with RAI and people typically needed they had to take 2 to 3 weeks off work. This means that THW was also considered to disadvantage those from lower socioeconomic groups, in whom a loss of earnings could adversely affect their quality of life. Additionally, with THW the person will become acutely hypothyroid and may experience mood changes such as anxiety and depression, lethargy and difficulty concentrating. It would therefore be important for them to be advised to avoid making important decisions during this time. This is particularly important for patients with pre-existing mental health problems. Therefore, the committee agreed that two groups of people would be disadvantaged with THW, those with a mental health disability (which is a protected characteristics under the equalities act) and those from a lower socioeconomic background.

The committee agreed that rhTSH should be recommended to other patient groups eligible for RAI because of the potential harm of THW. These groups included people with psychiatric conditions, cardiac conditions, older-age, chronic kidney disease and elevated falls risk who are not usually included in clinical trials.

The potential harm caused by THW was discussed. THW in preparation for RAI involves enforced seclusion for the patient, which may exacerbate certain psychiatric conditions, and this exacerbation may be increased by the hypothyroidism brought on by THW, which can adversely affect mood. Hypothyroidism may also increase risk of cardiac morbidity as a result of a reduction in cardiac output, which may be particularly harmful for those in heart failure. Meanwhile, older people may be more susceptible to the adverse events of hypothyroidism as a result of frailty reducing their ability to cope with such a stress. Whilst more severe kidney disease would be a contraindication to radioisotope therapy of any form, chronic kidney disease (CKD) per se is not an absolute contraindication and should be assessed on an individual patient basis. The committee were aware that in specialist centres radioiodine is given to patients on dialysis. However, CKD would be exacerbated by the adverse renal effects of hypothyroidism (decreased glomerular filtration rate) and so would be a contraindication for preparation with THW. Finally, hypothyroidism may increase the risk of falls and so it was agreed that people at risk of falls should also be viewed as being contraindicated to THW. The committee were aware this this would include off label use of rhTSH for some people as it is not licenced for people with advanced cancers and other metastatic disease. However, this was not believed to threaten the validity of the recommendation, because the licensing regulations permit rhTSH if THW poses a threat to the patient's well-being.

Given that the evidence showed relative benefits for rhTSH without any clear attendant harms, the committee agreed that a recommendation should be made for rhTSH to be used as the preparatory strategy for RAI in the patient groups aligned to those in the review evidence. These were patients that were in the 'lower stages' of disease, such as those without T4 disease or distant metastases. Some in the committee cautioned whether the review evidence base was truly representative of the 'low-stage' non-metastatic population. However, after discussion, the committee agreed that there was no evidence that the evidence-base was non-representative. While the committee thought this should be a strong 'offer' recommendation they noted that the cost-effectiveness evidence from original analysis (discussed in the following section on cost-effectiveness and resource use) found a cost per QALY of thyrotropin alfa between £20,000 and £30,000. As a result, the committee a made consider recommendation for this group because thyrotropin alfa was not shown be cost-effective over THW at the £20,000 threshold. However, they noted that offering thyrotropin alfa rather than THW is current practice.

The committee agreed that most people in the wider thyroid cancer population who were eligible for RAI (even if outside the 'lower stages' population and those who were not contraindicated for THW) should also be considered for rhTSH. They acknowledged there

was no evidence in the wider population; however, they agreed that similar mechanisms to those operating in the reviewed populations would be likely to effect similar relative outcomes in such a wider population. Because this was based on consensus, 'consider' was used instead of offer. The committee were aware this this would be off label use of rhTSH but noted that this is current practice to use rhTSH in most people with thyroid cancer, including those with distant metastases.

The committee also agreed that thyrotropin alfa enabled people to return to normal activities within 2 or 3 days of treatment, whereas THW is taken for 4 to 6 weeks before treatment with RAI and people typically needed to take at least 2 to 3 weeks off work. This meant that THW was also considered to disadvantage those from lower socioeconomic groups, in whom a loss of earnings could adversely affect their quality of life. Therefore, the committee made a 'be aware' recommendation to highlight that taking time off work can disadvantage those for whom a loss of earnings could adversely affect their quality of life.

Whilst agreeing on the general benefits of using rhTSH over THW, the committee also discussed the possible harms of rhTSH. Although it was agreed that there were fewer patient groups vulnerable to harm from this approach, it was also agreed that in patients with CNS metastases the harms of rhTSH may exceed the benefits, causing significant tumour flare. Therefore, in such patients, considerable care would need to be used if rhTSH were given. An additional recommendation was therefore made to alert clinicians to ensure rhTSH was used with caution in people who have brain or spinal metastases.

1.1.12.4 Cost effectiveness and resource use

Four studies with relevant comparison were included in the economic literature review. These all compared endogenous stimulation of TSH with thyroid hormone withdrawal versus exogenous stimulation of THS with recombinant human TSH (rhTSH).

Three studies used a Markov model to extrapolate costs and QoL based on an early randomized controlled trial from Pacini 2006. Pacini 2006 collected quality of life as SF-36 only on a single follow up after randomization, so the authors had to use extrapolations and assumptions to estimate QALYs as SF-6D utility scores for the duration of their analyses. All the three trials estimated important benefits in terms of quality of life and found rhTSH to be either cost effective at a threshold of £20,000 or £30,000, or to dominate TH withdrawal.

A further within-trial analysis was based on the latest ESTIMABL randomized controlled trial and had to rely less on extrapolation as quality of life were collected during several follow-ups both as EQ-5D and SF-36. The analysis estimated a lower QALY gain associated with rhTSH (using both EQ-5D or SF-6D utility scores) and concluded that rhTSH is unlikely to be cost effective at current price.

As the conclusions on cost effectiveness of rhTSH were found to be heavily dependent on the trial chosen to inform the health economics analysis, an original cost-utility analysis was conducted using a meta-analysis of all the trials included in the clinical review. These were three studies: Pacini 2006, HiLo and ESTIMABL. Values from the three trials were meta-analysed together to estimate difference in quality of life at point of ablation. A linear transformation was applied to the SF-6D utility curve reported by Borget 2015 and based on ESTIMABL to fit a new EQ-5D utility curve with the same trend and shape of the previous curve but reflecting the meta-analysed EQ-5D values instead. A quality of life stimulation model was developed using unit costs from UK national sources such NHS Reference Costs 2019-2020 and BNF. The committee were involved in the analysis and their view was integrated in the model either as new data or, where data was unavailable, through the inclusion of several sensitivity or threshold analysis.

The model was made fully probabilistic and two main probabilistic scenarios were presented to the Committee (see TSH Model Economic Report). In the first scenario, it was assumed that around 50% of the people starting withdrawal would switch from Levothyroxine T4 to

Liothyronine T3 before initiating it. This is usually practiced as withdrawal from T3 is expected to be shorter because T3 has a much faster rate of clearance. The percentage was indirectly estimated using average withdrawal period reported in the Estimabl trial. The probabilistic analysis based on this scenario found a cost per QALY equal to £27,315 and a probability of being cost effective at a threshold of £20,000 or £30,000 of, respectively, 18% and 59%. A second scenario was presented excluding those receiving T3. The rationale for this scenario was the significantly higher cost of T3 in the UK compared to the rest of the world, which has been the reason for a recent CMA court action for “unfair price abuse” against the manufacturer. The price has been steadily declining since the start of the investigation in 2019 and it is possible it will reach the original price of £4 in the future. The scenario analysis showed that, if this happens, the cost per QALY would become similar to the one of Scenario 2, just above the £30,000 threshold. In Scenario 2, due to the lowest pharmaceutical costs in the withdrawal group, the probabilistic cost per QALY increased to £32,330 and the probability of being cost effective at a threshold of £20,000 or £30,000 became, respectively, 7% and 43%. Sensitivity scenario analyses showed that the cost per QALY decreases if more weight is given to Pacini 2006 trial or a higher usage of T3 was assumed. A threshold analysis on the level of adherence was presented to the committee. Adherence was defined as the proportion of people withdrawing from thyroid hormone showing up at RAI appointment with a non-adequate level of TSH. If this occurs, it is assumed they would still receive rhTSH to reach the level of TSH to receive the treatment. The threshold analysis showed that at a 85% level of adherence in the withdrawal group, rhTSH reaches cost-effectiveness at a threshold of £20,000 in the first scenario. In scenario 2, cost effectiveness at £30,000 or £20,000 thresholds were achieved with an adherence level of, respectively, 95% and 75%.

Most of the members of the committee were generally unfamiliar with thyroid hormone withdrawal as rhTSH has been favoured in the UK for the last two decades. In general, they were aware that adherence tends to be lower in the thyroid hormone group as there are cases of people undergoing withdrawal who failed to reach the TSH level required from their clinical experience. Moreover, the committee were aware of the harm of withdrawal in some particularly vulnerable people, as hypothyroidism can severely affect people’s physical and mental health and hinder them from performing daily life tasks including working. This is particularly disadvantageous for people with low paid jobs or with zero-hour contract as they would remain without a stable income during the weeks of hypothyroidism further decreasing their quality of life. These equality considerations were raised during the discussion and considered very important by the Committee.

Furthermore, the Committee were aware that, due to local inefficiencies of the NHS, it is not uncommon for people on withdrawal to receive RAI later than intended, thus experiencing hypothyroidism for a longer time than the clinically optimal time used in the trials. A change in current practice towards an increased use of thyroid hormone withdrawal may further disrupt NHS providers and prolong waiting time for RAI, which may lead to more people developing persistent disease due to a late ablation of thyroid tissue. Finally, the committee highlighted the importance for the society of reducing radiation exposure, which can be achieved through the use of rhTSH as radioactive clearance is generally faster with rhTSH compared to withdrawal resulting in a lower dose absorbed by the body and blood. This, in turn, should reduce the number of new diagnoses of second malignancies and other health issues associated with radiation exposure. Following the discussion and as rhTSH was found to be potentially cost effective in the health economics analysis, a general consensus was reached to recommend rhTSH against thyroid hormone withdrawal.

1.1.12.5 Other factors the committee took into account

The equality considerations for this recommendation related to people with mental health issues and those from a lower socioeconomic background are discussed in section 1.1.12.3 benefits and harms.

1.1.13 Recommendations supported by this evidence review

This evidence review supports recommendations 1.3.10 to 1.3.13.

References

1. Ara R, Brazier J. Deriving an algorithm to convert the eight mean SF-36 dimension scores into a mean EQ-5D preference-based score from published studies (where patient level data are not available). Value in Health. 2008; 11(7):1131-1143

2. Barbaro D, Boni G. Radioiodine ablation of post-surgical thyroid remnants after preparation with recombinant human TSH: why, how and when. European Journal of Surgical Oncology. 2007; 33(5):535-540

3. Barbaro D, Boni G, Meucci G, Simi U, Lapi P, Orsini P et al. Radioiodine treatment with 30 mCi after recombinant human thyrotropin stimulation in thyroid cancer: effectiveness for postsurgical remnants ablation and possible role of iodine content in L-thyroxine in the outcome of ablation. Journal of Clinical Endocrinology and Metabolism. 2003; 88(9):4110-4115

4. Barbaro D, Boni G, Meucci G, Simi U, Lapi P, Orsini P et al. Recombinant human thyroid-stimulating hormone is effective for radioiodine ablation of post-surgical thyroid remnants. Nuclear Medicine Communications. 2006; 27(8):627-632

5. Blamey S, Barraclough B, Delbridge L, Mernagh P, Standfield L, Weston A. Using recombinant human thyroid-stimulating hormone for the diagnosis of recurrent thyroid cancer. ANZ Journal of Surgery. 2005; 75(1-2):10-20

6. BMJ Group and the Royal Pharmaceutical Society of Great Britain. British National Formulary. Available from: https://bnf.nice.org.uk/ Last accessed: 24/03/2022.

7. Borget I, Bonastre J, Catargi B, Deandreis D, Zerdoud S, Rusu D et al. Quality of life and cost-effectiveness assessment of radioiodine ablation strategies in patients with thyroid cancer: Results from the randomized phase iii estimabl trial. Journal of Clinical Oncology. 2015; 33(26):2885-2892

8. Brazier J, Usherwood T, Harper R, Thomas K. Deriving a preference-based single index from the UK SF-36 Health Survey. Journal of Clinical Epidemiology. 1998; 51(11):1115-1128

9. Campenni A, Pignata SA, Baldari S. Post-operative radioiodine therapy (RaIT) as adjuvant therapy in low-intermediate risk differentiated thyroid cancer. Clinical and Translational Imaging. 2018; 6(5):347-355

10. Chianelli M, Todino V, Graziano FM, Panunzi C, Pace D, Guglielmi R et al. Low-activity (2.0 GBq; 54 mCi) radioiodine post-surgical remnant ablation in thyroid cancer: comparison between hormone withdrawal and use of rhTSH in low-risk patients. European Journal of Endocrinology. 2009; 160(3):431-436

11. Curtis L, Burns A. Unit costs of health and social care 2020. Canterbury. Personal Social Services Research Unit University of Kent, 2020. Available from: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2020/

12. Dehbi HM, Mallick U, Wadsley J, Newbold K, Harmer C, Hackshaw A. Recurrence after low-dose radioiodine ablation and recombinant human thyroid-stimulating hormone for differentiated thyroid cancer (HiLo): long-term results of an open-label, non-inferiority randomised controlled trial. The Lancet Diabetes & Endocrinology. 2019; 7(1):44-51

13. Doi SAR, Woodhouse NJY. Ablation of the thyroid remnant and 131 I dose in differentiated thyroid cancer. Clinical Endocrinology. 2000; 52(6):765-773

14. Emmanouilidis N, Muller JA, Jager MD, Kaaden S, Helfritz FA, Guner Z et al. Surgery and radioablation therapy combined: introducing a 1-week-condensed procedure bonding total thyroidectomy and radioablation therapy with recombinant human TSH. European Journal of Endocrinology. 2009; 161(5):763-769

15. Emmanouilidis N, Schrem H, Winkler M, Klempnauer J, Scheumann GFW. Long-term results after treatment of very low-, low-, and high-risk thyroid cancers in a combined setting of thyroidectomy and radio ablation therapy in euthyroidism. International Journal of Endocrinology. 2013; 2013:769473

16. Hanscheid H, Lassmann M, Luster M, Thomas SR, Pacini F, Ceccarelli C et al. Iodine biokinetics and dosimetry in radioiodine therapy of thyroid cancer: procedures and results of a prospective international controlled study of ablation after rhTSH or hormone withdrawal. Journal of Nuclear Medicine. 2006; 47(4):648-654

17. Health and Social Care Information Centre. Prescription cost analysis, England [updated December 2021]. Available from: https://www.nhsbsa.nhs.uk/prescription-data/dispensing-data/prescription-cost-analysis-pca-data Last accessed:

18. Iakovou I, Goulis DG, Tsinaslanidou Z, Giannoula E, Katsikaki G, Konstantinidis I. Effect of recombinant human thyroid-stimulating hormone or levothyroxine withdrawal on salivary gland dysfunction after radioactive iodine administration for thyroid remnant ablation. Head and Neck. 2016; 38 (Suppl 1):E227-230

19. Iizuka Y, Katagiri T, Ogura K, Inoue M, Nakamura K, Mizowaki T. Comparison of thyroid hormone withdrawal and recombinant human thyroid-stimulating hormone administration for adjuvant therapy in patients with intermediate- to high-risk differentiated thyroid cancer. Annals of Nuclear Medicine. 2020; 34(10):736-741

20. Lamartina L, Durante C, Filetti S, Cooper DS. Low-risk differentiated thyroid cancer and radioiodine remnant ablation: a systematic review of the literature. Journal of Clinical Endocrinology and Metabolism. 2015; 100(5):1748-1761

21. Lee J, Yun MJ, Nam KH, Chung WY, Soh EY, Park CS. Quality of life and effectiveness comparisons of thyroxine withdrawal, triiodothyronine withdrawal, and recombinant thyroid-stimulating hormone administration for low-dose radioiodine remnant ablation of differentiated thyroid carcinoma. Thyroid. 2010; 20(2):173-179

22. Luster M, Felbinger R, Dietlein M, Reiners C. Thyroid hormone withdrawal in patients with differentiated thyroid carcinoma: A one hundred thirty-patient pilot survey on consequences of hypothyroidism and a pharmacoeconomic comparison to recombinant thyrotropin administration. Thyroid. 2005; 15(10):1147-1155

23. Ma C, Xie J, Liu W, Wang G, Zuo S, Wang X et al. Recombinant human thyrotropin (rhTSH) aided radioiodine treatment for residual or metastatic differentiated thyroid cancer. Cochrane Database of Systematic Reviews 2010, Issue 11. Art. No.: CD008302. DOI: 10.1002/14651858.CD008302.pub2.

24. Mallick U, Harmer C, Hackshaw A. The HiLo trial: a multicentre randomised trial of high- versus low-dose radioiodine, with or without recombinant human thyroid stimulating hormone, for remnant ablation after surgery for differentiated thyroid cancer. Clinical Oncology (Royal College of Radiologists). 2008; 20(5):325-326

25. Mallick U, Harmer C, Yap B, Wadsley J, Clarke S, Moss L et al. Ablation with low-dose radioiodine and thyrotropin alfa in thyroid cancer. New England Journal of Medicine. 2012; 366(18):1674-1685

26. Marturano I, Russo M, Spadaro A, Latina A, Malandrino P, Regalbuto C. Comparison of conventional L-thyroxine withdrawal and moderate hypothyroidism in preparation

for whole-body 131-I scan and thyroglobulin testing. Journal of Endocrinological Investigation. 2015; 38(9):1017-1022

27. Mernagh P, Campbell S, Dietlein M, Luster M, Mazzaferri E, Weston AR. Cost-effectiveness of using recombinant human TSH prior to radioiodine ablation for thyroid cancer, compared with treating patients in a hypothyroid state: the German perspective. European Journal of Endocrinology. 2006; 155(3):405-414

28. Mernagh P, Suebwongpat A, Silverberg J, Weston A. Cost-effectiveness of using recombinant human thyroid-stimulating hormone before radioiodine ablation for thyroid cancer: the Canadian perspective. Value in Health. 2010; 13(2):180-187

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

30. NHS England and NHS Improvement. National Cost Collection Data Publication 2019-2020. London. 2020. Available from: https://www.england.nhs.uk/wp-content/uploads/2021/06/National-Cost-Collection-2019-20-Report-FINAL.pdf

31. Nygaard B, Bastholt L, Bennedbaek FN, Klausen TW, Bentzen J. A placebo-controlled, blinded and randomised study on the effects of recombinant human thyrotropin on quality of life in the treatment of thyroid cancer. European Thyroid Journal. 2013; 2(3):195-202

32. Organisation for Economic Co-operation and Development (OECD). Purchasing power parities (PPP). 2021. Available from: http://www.oecd.org/std/ppp Last accessed: 24/03/2022.

33. Pacini F, Ladenson PW, Schlumberger M, Driedger A, Luster M, Kloos RT et al. Radioiodine ablation of thyroid remnants after preparation with recombinant human thyrotropin in differentiated thyroid carcinoma: results of an international, randomized, controlled study. Journal of Clinical Endocrinology and Metabolism. 2006; 91(3):926-932

34. Pacini F, Molinaro E, Castagna MG, Lippi F, Ceccarelli C, Agate L et al. Ablation of thyroid residues with 30 mCi (131)I: a comparison in thyroid cancer patients prepared with recombinant human TSH or thyroid hormone withdrawal. Journal of Clinical Endocrinology and Metabolism. 2002; 87(9):4063-4068

35. Pak K, Cheon GJ, Kang KW, Kim SJ, Kim IJ, Kim EE et al. The effectiveness of recombinant human thyroid-stimulating hormone versus thyroid hormone withdrawal prior to radioiodine remnant ablation in thyroid cancer: a meta-analysis of randomized controlled trials. Journal of Korean Medical Science. 2014; 29(6):811-817

36. Robbins RJ, Larson SM, Sinha N, Shaha A, Divgi C, Pentlow KS et al. A retrospective review of the effectiveness of recombinant human TSH as a preparation for radioiodine thyroid remnant ablation. Journal of Nuclear Medicine. 2002; 43(11):1482-1488

37. Schlumberger M, Catargi B, Borget I, Deandreis D, Zerdoud S, Bridji B et al. Strategies of radioiodine ablation in patients with low-risk thyroid cancer. New England Journal of Medicine. 2012; 366(18):1663-1673

38. Sohn SY, Jang HW, Cho YY, Kim SW, Chung JH. Economic evaluation of recombinant human thyroid stimulating hormone stimulation vs. Thyroid hormone withdrawal prior to radioiodine ablation for thyroid cancer: The korean perspective. Endocrinology and Metabolism. 2015; 30(4):531-542

39. Taieb D, Sebag F, Cherenko M, Baumstarck-Barrau K, Fortanier C, Farman-Ara B et al. Quality of life changes and clinical outcomes in thyroid cancer patients undergoing radioiodine remnant ablation (RRA) with recombinant human TSH (rhTSH): a randomized controlled study. Clinical Endocrinology. 2009; 71(1):115-123

40. Taieb D, Sebag F, Farman-Ara B, Portal T, Baumstarck-Barrau K, Fortanier C et al. Iodine biokinetics and radioiodine exposure after recombinant human thyrotropin-assisted remnant ablation in comparison with thyroid hormone withdrawal. Journal of Clinical Endocrinology and Metabolism. 2010; 95(7):3283-3290

41. Tu J, Wang S, Huo Z, Lin Y, Li X, Wang S. Recombinant human thyrotropin-aided versus thyroid hormone withdrawal-aided radioiodine treatment for differentiated thyroid cancer after total thyroidectomy: a meta-analysis. Radiotherapy and Oncology. 2014; 110(1):25-30

42. Vaiano A, Claudio Traino A, Boni G, Grosso M, Lazzeri P, Colato C et al. Comparison between remnant and red-marrow absorbed dose in thyroid cancer patients submitted to 131I ablative therapy after rh-TSH stimulation versus hypothyroidism induced by L-thyroxine withdrawal. Nuclear Medicine Communications. 2007; 28(3):215-223

43. Vallejo JA, Muros MA. Cost-effectiveness of using recombinant human thyroid-stimulating hormone before radioiodine ablation for thyroid cancer treatment in Spanish hospitals. Revista Espanola de Medicina Nuclear e Imagen Molecular. 2017; 36(6):362-370

44. van der Horst-Schrivers AN, Sluiter WJ, Muller Kobold AC, Wolffenbuttel BH, Plukker JT, Bisschop PH et al. Recombinant TSH stimulated remnant ablation therapy in thyroid cancer: the success rate depends on the definition of ablation success--an observational study. PLoS ONE [Electronic Resource]. 2015; 10(3):e0120184

45. Verburg FA, Flux G, Giovanella L, van Nostrand D, Muylle K, Luster M. Differentiated thyroid cancer patients potentially benefitting from postoperative I-131 therapy: a review of the literature of the past decade. European Journal of Nuclear Medicine and Molecular Imaging. 2020; 47(1):78-83

46. Waissi F, Kist JW, Lodewijk L, De Wit AG, Van Der Hage JA, Van Dalen T et al. Fast-track radioiodine ablation therapy after thyroidectomy reduces sick leave in patients with differentiated thyroid cancer (fasthyna trial). Clinical Nuclear Medicine. 2019; 44(4):272-275

47. Xu G, Wu T, Ge L, Li W. A systematic review of adjuvant interventions for radioiodine in patients with thyroid cancer. Oncology Research and Treatment. 2015; 38(7-8):368-372

48. Yoo J, Cosby R, Driedger A. Preparation with recombinant humanized thyroid-stimulating hormone before radioiodine ablation after thyroidectomy: a systematic review. Current Oncology. 2009; 16(5):23-31

Appendices

Appendix A – Review protocols

A.1 Review protocol for radioiodine ablation with withdrawal of levothyroxine to radioiodine ablation with thyrotropin alfa

Field	Content
PROSPERO registration number	CRD42020213225
Review title	Clinical and cost effectiveness of **radioactive iodine with withdrawal of thyroid hormone replacement versus radioactive iodine with thyrotropin alfa**, for people deemed suitable for RAI treatment who have had thyroidectomy for differentiated thyroid cancer.
Review question	What is the clinical and cost effectiveness of radioactive iodine with withdrawal of thyroid hormone replacement versus radioactive iodine with thyrotropin alfa?
Objective	To determine the best strategy of RAI ablation/treatment after surgery for differentiated thyroid cancer.
Searches	The following databases (from inception) will be searched: • Cochrane Central Register of Controlled Trials (CENTRAL) • Cochrane Database of Systematic Reviews (CDSR) • Embase • MEDLINE Searches will be restricted by:

	• English language • Human studies • Letters and comments are excluded. Other searches: • Inclusion lists of relevant systematic reviews will be checked by the reviewer. The searches may be re-run 6 weeks before final committee meeting and further studies retrieved for inclusion if relevant. The full search strategies will be published in the final review.
Condition or domain being studied	Thyroid cancer
Population	Inclusion: People aged 16 or over who have had thyroidectomy for differentiated thyroid cancer, and who are deemed suitable for RAI ablation/treatment. Exclusion: Children under 16
Intervention/Exposure/Test	• radioactive iodine ablation/treatment with prior withdrawal of thyroid hormone replacement • radioactive iodine ablation/treatment with prior preparation with thyrotropin alfa

Comparator/Reference standard/Confounding factors	• Each other • radioactive iodine ablation/treatment with neither of the above two uptake-stimulating strategies
Types of study to be included	• Systematic reviews • RCTs Non-randomised studies (any controlled designs, such as prospective/retrospective cohorts and case-control studies, with evidence of adjustment for biologically plausible confounders) will be included for one/both strata (ablation/treatment) *if* there are no RCTs in one/both strata.
Other exclusion criteria	Non-English language studies. Conference abstracts will be excluded as it is expected there will be sufficient full text published studies available.
Context	There is currently uncertainty about the best methods of providing RAI. in particular, the best method to ensure adequate iodine uptake to thyroid tissue is currently not established.
Primary outcomes (critical outcomes)	• mortality • quality of life (any validated scales) • local cancer progression (increase in size/number of tumours) • incidence of distant metastases • cancer recurrence • successful ablation • Second primary malignancy Longest available follow up in the studies.
Secondary outcomes (important outcomes)	None

Data extraction (selection and coding)	EndNote will be used for reference management, sifting, citations and bibliographies. All references identified by the searches and from other sources will be screened for inclusion. 10% of the abstracts will be reviewed by two reviewers, with any disagreements resolved by discussion or, if necessary, a third independent reviewer. The full text of potentially eligible studies will be retrieved and will be assessed in line with the criteria outlined above. A standardised form will be used to extract data from studies (see Developing NICE guidelines: the manual section 6.4). 10% of all evidence reviews are quality assured by a senior research fellow. This includes checking: • papers were included /excluded appropriately • a sample of the data extractions • correct methods are used to synthesise data • a sample of the risk of bias assessments Disagreements between the review authors over the risk of bias in particular studies will be resolved by discussion, with involvement of a third review author where necessary.
Risk of bias (quality) assessment	Risk of bias will be assessed using the appropriate checklist as described in Developing NICE guidelines: the manual. For Intervention reviews the following checklist will be used according to study design being assessed: • Systematic reviews: Risk of Bias in Systematic Reviews (ROBIS) • Randomised Controlled Trial: Cochrane RoB (2.0) • Non randomised study, including cohort studies: Cochrane ROBINS-I (if a lack of any RCTs necessitate dropping down to non-randomised studies) 10% of all evidence reviews are quality assured by a senior research fellow. This includes checking: • papers were included /excluded appropriately • a sample of the data extractions

	• correct methods are used to synthesise data • a sample of the risk of bias assessments Disagreements between the review authors over the risk of bias in particular studies will be resolved by discussion, with involvement of a third review author where necessary.
Strategy for data synthesis	Where possible, data will be meta-analysed. Pairwise meta-analyses will be performed using Cochrane Review Manager (RevMan5) to combine the data given in all studies for each of the outcomes stated above. A fixed effect meta-analysis, with weighted mean differences for continuous outcomes and risk ratios for binary outcomes will be used, and 95% confidence intervals will be calculated for each outcome. Heterogeneity between the studies in effect measures will be assessed using the I^2 statistic and visually inspected. We will consider an I^2 value greater than 50% indicative of substantial heterogeneity. Sensitivity analyses will be conducted based on pre-specified subgroups using stratified meta-analysis to explore the heterogeneity in effect estimates. If this does not explain the heterogeneity, the results will be presented using random-effects. GRADE pro will be used to assess the quality of each outcome, taking into account individual study quality and the meta-analysis results. The 4 main quality elements (risk of bias, indirectness, inconsistency and imprecision) will be appraised for each outcome. Publication bias is tested for when there are more than 5 studies for an outcome. Other bias will only be taken into consideration in the quality assessment if it is apparent. Where meta-analysis is not possible, data will be presented and quality assessed individually per outcome. If sufficient data is available to make a network of treatments, WinBUGS will be used for network meta-analysis.
Analysis of sub-groups	<u>Stratification (up-front stratification of analysis, NOT conditional on heterogeneity of prior meta-analysis)</u> • ablation vs treatment/therapy <u>Sub-grouping (conditional stratification if heterogeneity seen in initial unstratified meta-analysis)</u> If serious or very serious heterogeneity (I2>50%) is present within any stratum, sub-grouping will occur according to the following strategy: • Additional dietary restrictions vs no additional dietary restrictions • TSH levels normal (<30) or high (≥30)

	• Activity low (1Gb) vs higher (3-4 Gb)
Type and method of review	☒ Intervention ☐ Diagnostic ☐ Prognostic ☐ Qualitative ☐ Epidemiologic ☐ Service Delivery ☐ Other (please specify)
Language	English
Country	England
Named contact	**Named contact** National Guideline Centre **Organisational affiliation of the review** National Institute for Health and Care Excellence (NICE) and the National Guideline Centre
Review team members	From the National Guideline Centre: Carlos Sharpin, Guideline lead Mark Perry, Senior systematic reviewer Vimal Bedia, Systematic reviewer David Wonderling, Head of health economics Alfredo Mariani, Health economist

	Shama Mahammed, Health economist Lina Gulhane, Head of Information specialists
Funding sources/sponsor	This systematic review is being completed by the National Guideline Centre which receives funding from NICE.
Conflicts of interest	All guideline committee members and anyone who has direct input into NICE guidelines (including the evidence review team and expert witnesses) must declare any potential conflicts of interest in line with NICE's code of practice for declaring and dealing with conflicts of interest. Any relevant interests, or changes to interests, will also be declared publicly at the start of each guideline committee meeting. Before each meeting, any potential conflicts of interest will be considered by the guideline committee Chair and a senior member of the development team. Any decisions to exclude a person from all or part of a meeting will be documented. Any changes to a member's declaration of interests will be recorded in the minutes of the meeting. Declarations of interests will be published with the final guideline.
Collaborators	Development of this systematic review will be overseen by an advisory committee who will use the review to inform the development of evidence-based recommendations in line with section 3 of Developing NICE guidelines: the manual. Members of the guideline committee are available on the NICE website: https://www.nice.org.uk/guidance/indevelopment/gid-ng10150/documents
Other registration details	N/A
Reference/URL for published protocol	https://www.crd.york.ac.uk/PROSPERO/display_record.php?RecordID=213225
Dissemination plans	NICE may use a range of different methods to raise awareness of the guideline. These include standard approaches such as: • notifying registered stakeholders of publication • publicising the guideline through NICE's newsletter and alerts • issuing a press release or briefing as appropriate, posting news articles on the NICE website, using social media channels, and publicising the guideline within NICE.

Keywords	N/A
Details of existing review of same topic by same authors	N/A
Additional information	N/A
Details of final publication	www.nice.org.uk

A.2 Review protocol health economic evidence

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 2005, abstract-only studies and studies from non-OECD countries or the USA will also be excluded. Each remaining study will be assessed for applicability and methodological limitations using the NICE economic evaluation checklist which can be found in appendix H of Developing NICE guidelines: the manual (2014).[29] **Inclusion and exclusion criteria** • If a study is rated as both 'Directly applicable' and with 'Minor limitations', then it will be included in the guideline. A health economic evidence table will be completed, and it will be included in the health economic evidence profile. • If a study is rated as either 'Not applicable' or with 'Very serious limitations', then it will usually be excluded from the guideline. If it is excluded, then a health economic evidence table will not be completed, and it will not be included in the health economic evidence profile. • If a study is rated as 'Partially applicable', with 'Potentially serious limitations' or both then there is discretion over whether it should be included. **Where there is discretion**

The health economist will make a decision based on the relative applicability and quality of the available evidence for that question, in discussion with the guideline committee if required. The ultimate aim is to include health economic studies that are helpful for decision-making in the context of the guideline and the current NHS setting. If several studies are considered of sufficiently high applicability and methodological quality that they could all be included, then the health economist, in discussion with the committee if required, may decide to include only the most applicable studies and to selectively exclude the remaining studies. All studies excluded on the basis of applicability or methodological limitations will be listed with explanation in the excluded health economic studies appendix below.

The health economist will be guided by the following hierarchies.

Setting:

- UK NHS (most applicable).
- OECD countries with predominantly public health insurance systems (for example, France, Germany, Sweden).
- OECD countries with predominantly private health insurance systems (for example, Switzerland).
- Studies set in non-OECD countries or in the USA will be excluded before being assessed for applicability and methodological limitations.

Health economic study type:

- Cost–utility analysis (most applicable).
- Other type of full economic evaluation (cost–benefit analysis, cost-effectiveness analysis, cost–consequences analysis).
- Comparative cost analysis.
- Non-comparative cost analyses including cost-of-illness studies will be excluded before being assessed for applicability and methodological limitations.

Year of analysis:

- The more recent the study, the more applicable it will be.
- Studies published in 2005 or later but that depend on unit costs and resource data entirely or predominantly from before 2005 will be rated as ‘Not applicable’.
- Studies published before 2005 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 these reviews are detailed below and complied with the methodology outlined in Developing NICE guidelines: the manual, 2014 (updated 2020) https://www.nice.org.uk/process/pmg20/chapter/identifying-the-evidence-literature-searching-and-evidence-submission.

For more information, please see the Methodology review published as part of the accompanying documents for this guideline.

Clinical literature search strategy

This literature search strategy was used for this review:

- What is the clinical and cost effectiveness of radioactive iodine with thyrotropin alfa versus radioactive iodine with withdrawal of thyroid hormone replacement?

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 9: Database parameters, filters and limits applied

Database	Dates searched	Search filters and limits applied
Medline (OVID)	1946 – 13 January 2022	Randomised controlled trials Systematic review studies Observational studies Exclusions (animal studies, letters, comments, editorials, case studies/reports, children) English language
Embase (OVID)	1974 – 13 January 2022	Randomised controlled trials Systematic review studies Observational studies Exclusions (animal studies, letters, comments, editorials, case studies/reports, conference abstracts, children) English language
The Cochrane Library (Wiley)	Cochrane Database of Systematic Reviews to Issue 12 of 12, December 2021 Cochrane Central Register of Controlled Trials to Issue 12 of 12, December 2021	Exclusions (clinical trials, conference abstracts)

Database	Dates searched	Search filters and limits applied
Epistemonikos (The Epistemonikos Foundation)	Inception – 13 January 2022	Systematic review Exclusions (Cochrane reviews) English language

1

2 **Medline (Ovid) search terms**

1.	exp Thyroid Neoplasms/
2.	(thyroid and (cancer* or carcinom* or microcarcinoma* or tumo?r* or neoplasm* or metast* or adenoma* or adenocarcinom* or node* or nodul* or nodal or lump* or papillar* or swollen or swell* or follicul* or lymphoma* or anaplastic or sarcoma* or medullar* or cyst* or malignan*)).ti,ab.
3.	DTC.ti,ab.
4.	((papillar* or follicul* or medullar* or anaplastic) adj2 (cancer* or carcinom* or tumo?r* or neoplasm* or metast* or adenoma* or adenocarcinom* or nodul* or node* or lump* or lymphoma*)).ti,ab.
5.	or/1-4
6.	letter/
7.	editorial/
8.	news/
9.	exp historical article/
10.	Anecdotes as Topic/
11.	comment/
12.	case report/
13.	(letter or comment*).ti.
14.	or/6-13
15.	randomized controlled trial/ or random*.ti,ab.
16.	14 not 15
17.	animals/ not humans/
18.	exp Animals, Laboratory/
19.	exp Animal Experimentation/
20.	exp Models, Animal/
21.	exp Rodentia/
22.	(rat or rats or mouse or mice or rodent*).ti.
23.	or/16-22
24.	5 not 23
25.	limit 24 to english language
26.	exp radiotherapy/
27.	radiotherapy dosage/
28.	Iodine Radioisotopes/
29.	radioiodine.ti,ab.
30.	(iodi?e adj2 (radio* or isotope*)).ti,ab.
31.	(iodi?e 131 or 131-I or I-131).ti,ab.
32.	remnant ablation.ti,ab.

33.	(iodi?e adj2 (ablation or treatment* or therap* or medic* or procedure* or intervention*)).ti,ab.
34.	(RAA or RRA or RAI).ti,ab.
35.	or/26-34
36.	25 and 35
37.	randomized controlled trial.pt.
38.	controlled clinical trial.pt.
39.	randomi#ed.ab.
40.	placebo.ab.
41.	randomly.ab.
42.	clinical trials as topic.sh.
43.	trial.ti.
44.	or/37-43
45.	Meta-Analysis/
46.	Meta-Analysis as Topic/
47.	(meta analy* or metanaly* or metaanaly* or meta regression).ti,ab.
48.	((systematic* or evidence*) adj3 (review* or overview*)).ti,ab.
49.	(reference list* or bibliograph* or hand search* or manual search* or relevant journals).ab.
50.	(search strategy or search criteria or systematic search or study selection or data extraction).ab.
51.	(search* adj4 literature).ab.
52.	(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.
53.	cochrane.jw.
54.	((multiple treatment* or indirect or mixed) adj2 comparison*).ti,ab.
55.	or/45-54
56.	Epidemiologic studies/
57.	Observational study/
58.	exp Cohort studies/
59.	(cohort adj (study or studies or analys* or data)).ti,ab.
60.	((follow up or observational or uncontrolled or non randomi#ed or epidemiologic*) adj (study or studies or data)).ti,ab.
61.	((longitudinal or retrospective or prospective or cross sectional) and (study or studies or review or analys* or cohort* or data)).ti,ab.
62.	Controlled Before-After Studies/
63.	Historically Controlled Study/
64.	Interrupted Time Series Analysis/
65.	(before adj2 after adj2 (study or studies or data)).ti,ab.
66.	exp case control study/
67.	case control*.ti,ab.
68.	Cross-sectional studies/
69.	(cross sectional and (study or studies or review or analys* or cohort* or data)).ti,ab.
70.	or/57-70
71.	36 and (44 or 55 or 70)

1 **Embase (Ovid) search terms**

1.	exp Thyroid Cancer/

2.	(thyroid adj3 (cancer* or carcinom* or microcarcinoma* or tumo?r* or neoplasm* or metast* or adenoma* or adenocarcinom* or node* or nodul* or nodal or lump* or papillar* or swollen or swell* or anaplastic or sarcoma* or cyst* or malignan*)).ti,ab.
3.	DTC.ti,ab.
4.	((papillar* or anaplastic) adj2 (cancer* or carcinom* or tumo?r* or neoplasm* or metast* or adenoma* or adenocarcinom* or nodul* or node* or lump*)).ti,ab.
5.	or/1-4
6.	letter.pt. or letter/
7.	note.pt.
8.	editorial.pt.
9.	case report/ or case study/
10.	(letter or comment*).ti.
11.	(conference abstract or conference paper).pt.
12.	or/6-11
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 or rodent*).ti.
22.	or/14-21
23.	5 not 22
24.	limit 23 to english language
25.	(exp child/ or exp pediatrics/) not (exp adult/ or exp adolescent/)
26.	24 not 25
27.	exp radiotherapy/
28.	radiotherapy dosage/
29.	radioactive iodine/
30.	radioiodine.ti,ab.
31.	(iodi?e adj2 (radio* or isotope*)).ti,ab.
32.	iodine 131/
33.	(iodi?e 131 or 131-I or I-131).ti,ab.
34.	remnant ablation.ti,ab.
35.	(iodi?e adj2 (ablation or treatment* or therap* or medic* or procedure* or intervention*)).ti,ab.
36.	(RAA or RRA or RAI).ti,ab.
37.	or/27-36
38.	26 and 37
39.	random*.ti,ab.
40.	factorial*.ti,ab.
41.	(crossover* or cross over*).ti,ab.
42.	((doubl* or singl*) adj blind*).ti,ab.
43.	(assign* or allocat* or volunteer* or placebo*).ti,ab.
44.	crossover procedure/
45.	single blind procedure/

46.	randomized controlled trial/
47.	double blind procedure/
48.	or/39-47
49.	systematic review/
50.	Meta-Analysis/
51.	(meta analy* or metanaly* or metaanaly* or meta regression).ti,ab.
52.	((systematic* or evidence*) adj3 (review* or overview*)).ti,ab.
53.	(reference list* or bibliograph* or hand search* or manual search* or relevant journals).ab.
54.	(search strategy or search criteria or systematic search or study selection or data extraction).ab.
55.	(search* adj4 literature).ab.
56.	(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.
57.	cochrane.jw.
58.	((multiple treatment* or indirect or mixed) adj2 comparison*).ti,ab.
59.	or/49-58
60.	Clinical study/
61.	Observational study/
62.	family study/
63.	longitudinal study/
64.	retrospective study/
65.	prospective study/
66.	cohort analysis/
67.	follow-up/
68.	cohort*.ti,ab.
69.	67 and 68
70.	(cohort adj (study or studies or analys* or data)).ti,ab.
71.	((follow up or observational or uncontrolled or non randomi#ed or epidemiologic*) adj (study or studies or data)).ti,ab.
72.	((longitudinal or retrospective or prospective) and (study or studies or review or analys* or cohort* or data)).ti,ab.
73.	(before adj2 after adj2 (study or studies or data)).ti,ab.
74.	exp case control study/
75.	case control*.ti,ab.
76.	cross-sectional study/
77.	(cross sectional and (study or studies or review or analys* or cohort* or data)).ti,ab.
78.	or/60-66,69-77
79.	38 and (48 or 59 or 78)

Cochrane Library (Wiley) search terms

#1.	MeSH descriptor: [Thyroid Neoplasms] explode all trees
#2.	(thyroid near/3 (cancer* or carcinom* or microcarcinoma* or tumo?r* or neoplasm* or metast* or adenoma* or adenocarcinom* or node* or nodul* or nodal or lump* or papillar* or swollen or swell* or anaplastic or sarcoma* or cyst* or malignan*)):ti,ab
#3.	DTC:ti,ab
#4.	((papillar* or anaplastic) near/2 (cancer* or carcinom* or tumo?r* or neoplasm* or metast* or adenoma* or adenocarcinom* or nodul* or node* or lump*)):ti,ab
#5.	#1 or #2 or #3 or #4

#6.	conference:pt or (clinicaltrials or trialsearch):so
#7.	#5 not #6
#8.	MeSH descriptor: [Iodine Radioisotopes] explode all trees
#9.	MeSH descriptor: [Radiotherapy] explode all trees
#10.	MeSH descriptor: [Radiotherapy Dosage] this term only
#11.	radioiodine:ti,ab
#12.	((iodi?e) near/2 (radio* or isotope*)):ti,ab
#13.	(iodi?e-131 or I-131):ti,ab
#14.	remnant ablation:ti,ab
#15.	((iodi?e) near/2 (ablation or treatment* or therap* or medic* or procedure* or intervention*)):ti,ab
#16.	(RAA or RRA or RAI):ti,ab
#17.	#8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16
#18.	#7 and #17

Epistemonikos search terms

1.	(title:(remnant ablation OR RAI OR RRA OR RAA) OR abstract:(remnant ablation OR RAI OR RRA OR RAA)) OR (title:(thyroid AND (iodine OR iodide)) OR abstract:(thyroid AND (iodine OR iodide)))

Health Economics literature search strategy

Health economic evidence was identified by conducting searches using terms for a broad Thyroid Cancer population. The following databases were searched: NHS Economic Evaluation Database (NHS EED - this ceased to be updated after 31st March 2015), Health Technology Assessment database (HTA - this ceased to be updated from 31st March 2018) and The International Network of Agencies for Health Technology Assessment (INAHTA). Searches for recent evidence were run on Medline and Embase from 2014 onwards for health economics, and all years for quality-of-life studies.

Table 2: Database parameters, filters and limits applied

Database	Dates searched	Search filters and limits applied
Medline (OVID)	Health Economics 1 January 2014 – 16 December 2021	Health economics studies Quality of life studies
	Quality of Life 1946 – 16 December 2021	Exclusions (animal studies, letters, comments, editorials, case studies/reports, conference abstracts) English language
Embase (OVID)	Health Economics 1 January 2014 – 16 December 2021	Health economics studies Quality of life studies
	Quality of Life 1974 – 16 December 2021	Exclusions (animal studies, letters, comments, editorials, case studies/reports, conference abstracts) English language

Database	Dates searched	Search filters and limits applied
NHS Economic Evaluation Database (NHS EED) (Centre for Research and Dissemination - CRD)	Inception –31st March 2015	
Health Technology Assessment Database (HTA) (Centre for Research and Dissemination – CRD)	Inception – 31st March 2018	
The International Network of Agencies for Health Technology Assessment (INAHTA)	Inception - 16 December 2021	English language

1 **Medline (Ovid) search terms**

1.	exp Thyroid Neoplasms/
2.	(thyroid adj4 (cancer* or carcinom* or tumo?r* or neoplasm* or metast* or adenoma* or adenocarcinom* or nod* or lump* or papillar* or follicul* or lymphoma* or anaplastic)).ti,ab.
3.	((papillar* or follicul* or medullary or anaplastic) adj4 (cancer* or carcinom* or tumo?r* or neoplasm* or metast* or adenoma* or adenocarcinom* or nod* or lump* or lymphoma*)).ti,ab.
4.	or/1-3
5.	letter/
6.	editorial/
7.	news/
8.	exp historical article/
9.	Anecdotes as Topic/
10.	comment/
11.	case report/
12.	(letter or comment*).ti.
13.	or/5-12
14.	randomized controlled trial/ or random*.ti,ab.
15.	13 not 14
16.	animals/ not humans/
17.	exp Animals, Laboratory/
18.	exp Animal Experimentation/
19.	exp Models, Animal/
20.	exp Rodentia/
21.	(rat or rats or mouse or mice).ti.
22.	or/15-21
23.	4 not 22
24.	limit 23 to english language
25.	economics/
26.	value of life/
27.	exp "costs and cost analysis"/
28.	exp Economics, Hospital/
29.	exp Economics, medical/
30.	Economics, nursing/

31.	economics, pharmaceutical/
32.	exp "Fees and Charges"/
33.	exp budgets/
34.	budget*.ti,ab.
35.	cost*.ti.
36.	(economic* or pharmaco?economic*).ti.
37.	(price* or pricing*).ti,ab.
38.	(cost* adj2 (effectiv* or utilit* or benefit* or minimi* or unit* or estimat* or variable*)).ab.
39.	(financ* or fee or fees).ti,ab.
40.	(value adj2 (money or monetary)).ti,ab.
41.	or/25-40
42.	24 and 41
43.	quality-adjusted life years/
44.	sickness impact profile/
45.	(quality adj2 (wellbeing or well being)).ti,ab.
46.	sickness impact profile.ti,ab.
47.	disability adjusted life.ti,ab.
48.	(qal* or qtime* or qwb* or daly*).ti,ab.
49.	(euroqol* or eq5d* or eq 5*).ti,ab.
50.	(qol* or hql* or hqol* or h qol* or hrqol* or hr qol*).ti,ab.
51.	(health utility* or utility score* or disutilit* or utility value*).ti,ab.
52.	(hui or hui1 or hui2 or hui3).ti,ab.
53.	(health* year* equivalent* or hye or hyes).ti,ab.
54.	discrete choice*.ti,ab.
55.	rosser.ti,ab.
56.	(willingness to pay or time tradeoff or time trade off or tto or standard gamble*).ti,ab.
57.	(sf36* or sf 36* or short form 36* or shortform 36* or shortform36*).ti,ab.
58.	(sf20 or sf 20 or short form 20 or shortform 20 or shortform20).ti,ab.
59.	(sf12* or sf 12* or short form 12* or shortform 12* or shortform12*).ti,ab.
60.	(sf8* or sf 8* or short form 8* or shortform 8* or shortform8*).ti,ab.
61.	(sf6* or sf 6* or short form 6* or shortform 6* or shortform6*).ti,ab.
62.	or/52-70
63.	24 and 62

1 **Embase (Ovid) search terms**

1.	exp Thyroid Cancer/
2.	(thyroid adj4 (cancer* or carcinom* or tumo?r* or neoplasm* or metast* or adenoma* or adenocarcinom* or nod* or lump* or papillar* or follicul* or lymphoma* or anaplastic)).ti,ab.
3.	((papillar* or follicul* or medullary or anaplastic) adj4 (cancer* or carcinom* or tumo?r* or neoplasm* or metast* or adenoma* or adenocarcinom* or nod* or lump* or lymphoma*)).ti,ab.
4.	or/1-3
5.	letter.pt. or letter/
6.	note.pt.
7.	editorial.pt.
8.	case report/ or case study/
9.	(letter or comment*).ti.

10.	or/5-9
11.	randomized controlled trial/ or random*.ti,ab.
12.	10 not 11
13.	animal/ not human/
14.	nonhuman/
15.	exp Animal Experiment/
16.	exp Experimental Animal/
17.	animal model/
18.	exp Rodent/
19.	(rat or rats or mouse or mice).ti.
20.	or/12-19
21.	4 not 20
22.	limit 21 to english language
23.	health economics/
24.	exp economic evaluation/
25.	exp health care cost/
26.	exp fee/
27.	budget/
28.	funding/
29.	budget*.ti,ab.
30.	cost*.ti.
31.	(economic* or pharmaco?economic*).ti.
32.	(price* or pricing*).ti,ab.
33.	(cost* adj2 (effectiv* or utilit* or benefit* or minimi* or unit* or estimat* or variable*)).ab.
34.	(financ* or fee or fees).ti,ab.
35.	(value adj2 (money or monetary)).ti,ab.
36.	or/23-35
37.	22 and 36
38.	quality-adjusted life years/
39.	"quality of life index"/
40.	short form 12/ or short form 20/ or short form 36/ or short form 8/
41.	sickness impact profile/
42.	(quality adj2 (wellbeing or well being)).ti,ab.
43.	sickness impact profile.ti,ab.
44.	disability adjusted life.ti,ab.
45.	(qal* or qtime* or qwb* or daly*).ti,ab.
46.	(euroqol* or eq5d* or eq 5*).ti,ab.
47.	(qol* or hql* or hqol* or h qol* or hrqol* or hr qol*).ti,ab.
48.	(health utility* or utility score* or disutilit* or utility value*).ti,ab.
49.	(hui or hui1 or hui2 or hui3).ti,ab.
50.	(health* year* equivalent* or hye or hyes).ti,ab.
51.	discrete choice*.ti,ab.
52.	rosser.ti,ab.
53.	(willingness to pay or time tradeoff or time trade off or tto or standard gamble*).ti,ab.
54.	(sf36* or sf 36* or short form 36* or shortform 36* or shortform36*).ti,ab.
55.	(sf20 or sf 20 or short form 20 or shortform 20 or shortform20).ti,ab.
56.	(sf12* or sf 12* or short form 12* or shortform 12* or shortform12*).ti,ab.

57.	(sf8* or sf 8* or short form 8* or shortform 8* or shortform8*).ti,ab.
58.	(sf6* or sf 6* or short form 6* or shortform 6* or shortform6*).ti,ab.
59.	or/37-58
60.	22 and 59

NHS EED and HTA (CRD) search terms

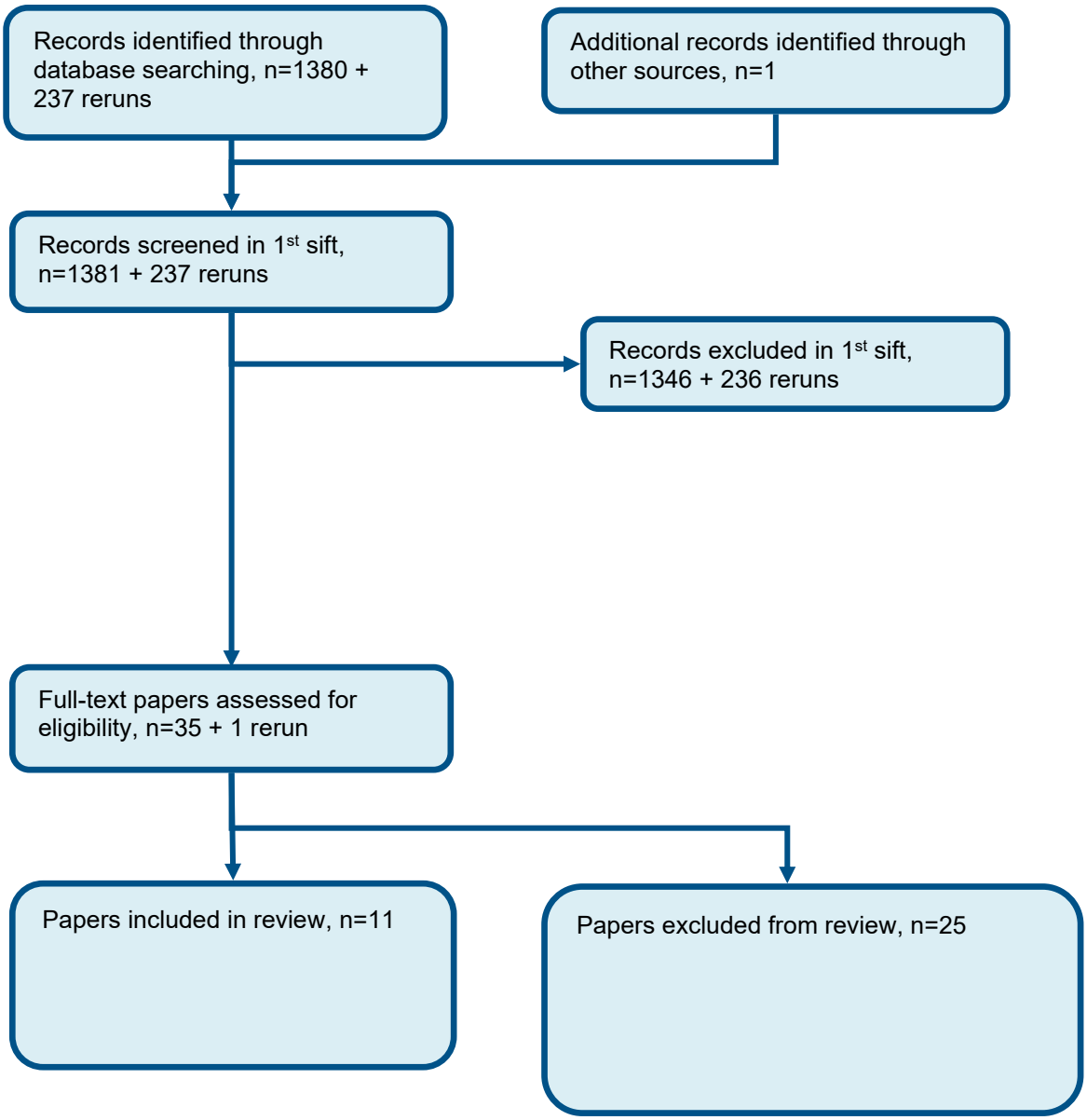
#1.	MeSH DESCRIPTOR Thyroid Neoplasms EXPLODE ALL TREES
#2.	((thyroid NEAR4 (cancer* or carcinom* or tumour* or tumor* or neoplasm* or metast* or adenoma* or adenocarcinom* or nod* or lump* or papillar* or follicul* or lymphoma* or anaplastic)))
#3.	(((papillar* or follicul* or medullary or anaplastic) NEAR4 (cancer* or carcinom* or tumour* or tumor* or neoplasm* or metast* or adenoma* or adenocarcinom* or nod* or lump* or lymphoma*)))
#4.	#1 OR #2 OR #3

INHATA search terms

1.	(Thyroid Neoplasms)[mh] OR (thyroid neoplasms) AND (thyroid cancers)

Appendix C – Effectiveness evidence study selection

Figure 4: Flow chart of clinical study selection for the review of radioactive iodine with or without preparation with thyrotropin alfa



Appendix D – Effectiveness evidence

Study	Chianelli 2009[10]
Study type	RCT (Patient randomised. Parallel)
Number of studies (number of participants)	(n=42)
Countries and setting	Conducted in Italy
Line of therapy	Not applicable
Duration of study	Intervention + follow up: 12 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Ablation
Subgroup analysis within study	Not applicable
Inclusion criteria	All patients had papillary cancer or minimally invasive follicular cancer, with a tumour node metastases stage pT1, larger than 1 cm or less than 1 cm if in the presence of multiple foci and could be considered patients at low risk of recurrence (stage I. tumour node metastases (TNM. staging according to AJCC 2002)) (9). No patient had positive cervical lymph nodes at the time of treatment as evaluated by US.
Exclusion criteria	Patients with positive Tg auto-antibodies were excluded from the study
Recruitment/selection of patients	All patients underwent total thyroidectomy or near-total thyroidectomy and, after surgery, began treatment with a TSH suppressive dose of L-T4. All patients adhered to a low-iodine diet for2weeks before receiving 131I.
Age, gender and ethnicity	Age - Mean (SD): Withdrawal: 48±9.9. rhTSH: 46.1±12.3. Gender (M:F): 9/33. Ethnicity:
Further population details	1. Activity level: Activity low (1Gb) ((2.0 GBq. 54 mCi)). 2. Diet: Additional dietary restrictions (low iodine diet for 2 weeks prior to 131I). 3. TSH levels: High (>30) (Hypothyroid: 77.9±17.1. rhTSH: 91.00±9.8).
Indirectness of population	No indirectness
Interventions	(n=21) Intervention 1: Radioactive iodine ablation - with prior withdrawal of thyroid hormone replacement. Twenty-one patients (age, 28–71 years. 16 females and 5 males) were treated with 131I in the hypothyroid state. L-T4 was stopped for 37 days. from the 3rd to 22nd day after L-T4 withdrawal patients were treated with T3. Patients received 131I (2.02±0.22 GBq. 54.6± 5.9 mCi. mean±S.D.) 42–180 days after surgery. On the day of

	administration of 131I, TSH, Tg, and TgAb were measured. L-T4 was then given again the day after administration of 131I.. Duration 37 days withdrawal up to treatment. Concurrent medication/care: NA. Indirectness: No indirectness (n=21) Intervention 2: Radioactive iodine ablation - with prior preparation with thyrotropin alfa. Twenty-one patients (age, 20–67 years. 17 females and 4 males) were treated with 131I following the administration of rhTSH (Thyrogen. Genzyme Corp, Cambridge, MA, USA): the therapeutic activity of 131I (1.97±0.18 GBq. 53.2±4.9 mCi. mean±S.D.) was administered 24 h after the last injection of rhTSH (0.9 mg i.m. for two consecutive days). L-T4 was never stopped during treatment. The time between thyroidectomy and 131I treatment was 42–180 days. Serum samples of TSH, FT4,FT3, Tg and anti-Tg antibodies were taken the day before the first administration of rhTSH. Serum samples for TSH, Tg and TgAb were also taken 3 days after the last administration of rhTSH. Levels of Tg (functional sensitivity: 0.7 ng/ml) were determined with a commercially available IRMA (Thyroglobuline IRMA. CIS-BIO, France). Serum levels of TSH (normal range 0.2–4.0, upper detection limit: 100 mIU/ml), free triiodothyronine (FT3, normal range 2.2–5.0 pg=ml), thyroxine (FT4, normal range 8.0–18.5 pg=ml) and anti-thyroglobulin antibodies (TgAb, normal range 0.0–70.0 IU=ml) were determined with commercially available radioimmunological assay kits (Radim, Pomezia, Italy). Urinary iodine excretion was measured to exclude contamination from stable iodine, using a colorimetric method (CellTech, Torino, Italy). . Duration two consecutive days of rhTSH prior to radioiodine treatment. Concurrent medication/care: NA. Indirectness: No indirectness
Funding	No funding

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: Withdrawal versus RHTSH

Protocol outcome 1: Successful ablation
- Actual outcome for Ablation: Ablation (Tg<1 ng/ml) at 6 months post treatment. Group 1: 18/20, Group 2: 17/20. Comments:
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: 1, Reason: not specified. Group 2 Number missing: 1, Reason: not specified

- Actual outcome for Ablation: Ablation (no visible uptake) at 6 months post treatment. Group 1: 20/21, Group 2: 19/21. Comments:
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: 0. Group 2 Number missing: 0

Protocol outcomes not reported by the study	Quality of life . Mortality . Local cancer progression . Incidence of distant metastases . Cancer recurrence

Study	Emmanouilidis 2009[14]
Study type	RCT (Patient randomised. Parallel)
Number of studies (number of participants)	(n=25)
Countries and setting	Conducted in Germany. Setting: Medical centre, Hanover Germany
Line of therapy	Not applicable
Duration of study	Intervention + follow up: 12 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Treatment / therapy
Subgroup analysis within study	Not applicable
Inclusion criteria	Patients with differentiated thyroid cancer and having a thyroidectomy and received a K1a/b central lymphadenectomy.
Exclusion criteria	not specified
Recruitment/selection of patients	cohort of patients with a diagnosis of DTC or from patients that were thyroidectomized due to multinodular struma and who had a coincidental histology of DTC
Age, gender and ethnicity	Age - Mean (SD): rhTSH: 45.2±16.5. Withdrawal: 54.8±12.8. Gender (M:F): 7/18. Ethnicity:
Further population details	1. Activity level: Activity higher (3 - 4 Gb) (3700MBq orally). 2. Diet: No additional dietary restrictions 3. TSH levels: Not stated / Unclear
Indirectness of population	No indirectness
Interventions	(n=13) Intervention 1: Radioactive iodine ablation - with prior preparation with thyrotropin alfa. RhTSH participants received their first RAT on first hospitalization. rhTSH (Thyrogen, Genzyme, Cambridge, MA, USA) with a biological potency of 10 U/mg of protein was used according to the manufacturer's instructions. Each vial containing 0.9 mg of rhTSH-alfa was dissolved in 1.2 ml of water for injection and administered by the i.m. route to the gluteal region 48 and 24 h before RAT.. Duration 24 - 48 hours before radioiodine ablation therapy. Concurrent medication/care: After iodine uptake was confirmed by neck scan with 100 MBq 131I, the ablative activity of 3700 MBq 131I was administered orally.. Indirectness: No indirectness (n=12) Intervention 2: Radioactive iodine ablation - with prior withdrawal of thyroid hormone replacement.. patients in L-T4 abstinence group were discharged from the surgery ward and, while in a state of distinctive hypothyroidism, were re-hospitalized for the first RAT within 4–6 weeks after thyroidectomy. Duration 4 - 6 weeks. Concurrent medication/care: After iodine uptake was confirmed by neck scan with 100 MBq 131I, the

	ablative activity of 3700 MBq 131I was administered orally.. Indirectness: No indirectness
Funding	No funding

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: RHTSH versus Withdrawal

Protocol outcome 1: Cancer recurrence
- Actual outcome for Ablation: Suspected tumour recurrence at day 96 - 131 post treatment. Group 1: 5/13, Group 2: 5/12. Comments: Additional RAT due to suspected tumour recurrence was conducted for three patients in rhTSH receivers and for four patients in L-T4 abstinence.
US by itself did not lead to additional RAT, whereas for two patients in rhTSH receivers and one patient in L-T4 abstinence a positive diagnostic scan lead to suspicion for tumour recurrence and thus were followed up by an additional ablative activity of 3700 MBq 131iodine, despite a negative US examination.
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - High, Measurement - High, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 0. Group 2 Number missing: 0

Protocol outcome 2: Successful ablation
- Actual outcome for Ablation: Thyroglobulin levels at final follow up . Group 1: mean 0.1 g/l (SD 0.27). n=13, Group 2: mean 0.28 g/l (SD 0.65). n=12. Comments: Baseline Tg rhTSH: 8.02 ng/l (16.47)
Baseline Tg Withdrawal: 8.26g/l (11.18)

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: 0. Group 2 Number missing: 0

Protocol outcomes not reported by the study	Quality of life . Mortality . Local cancer progression . Incidence of distant metastases

Study	Emmanouilidis 2013[15]
Study type	RCT (Patient randomised. Parallel)
Number of studies (number of participants)	(n=44)
Countries and setting	Conducted in Germany. Setting: Medical Centre, Hanover Germany
Line of therapy	Not applicable
Duration of study	Intervention + follow up: 4.5 years approximately
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Ablation
Subgroup analysis within study	Not applicable
Inclusion criteria	Patients with differentiated thyroid cancer
Exclusion criteria	Not specified
Recruitment/selection of patients	Patients with differentiated thyroid cancer awaiting radioiodine ablation therapy.
Age, gender and ethnicity	Age - Median (range): rhTSH: 50 (17-66). Withdrawal: 58 (30-73). Gender (M:F): 11/33. Ethnicity:
Further population details	1. Activity level: Activity higher (3 - 4 Gb) (3700MBq). 2. Diet: No additional dietary restrictions 3. TSH levels: Not stated / Unclear
Indirectness of population	No indirectness
Interventions	(n=24) Intervention 1: Radioactive iodine ablation - with prior preparation with thyrotropin alfa. RhTSH patients received their first RAT on first hospitalization. RhTSH (Tyrogen, Genzyme, Cambridge, Mass.) with a biological potency of 10 U/mg of protein was used according to the manufacturer's instructions. Each vial containing 0.9mg of rhTSH-alpha was dissolved in 1.2mL of water for injection and administered by the i.m. route to the gluteal region 48 h and 24 h before RAT. Duration 24 - 48h before radioiodine ablation therapy. Concurrent medication/care: After iodine uptake was confirmed by neck scan with 100Milli-Becquerel (MBq) 131I, the ablative activity of 3700MBq 131Iwas administered orally. (n=20) Intervention 2: Radioactive iodine ablation - with prior withdrawal of thyroid hormone replacement. patients in the L-T4 withdrawal group were discharged from hospital and readmitted for the first RAT within 4–6weeks after thyroidectomy while in a state of distinctive hypothyroidism. Duration 4 - 6 prior to radioiodine ablation therapy. Concurrent medication/care: After iodine uptake was confirmed by neck scan with 100Milli-Becquerel (MBq) 131I, the ablative activity of 3700MBq 131Iwas administered orally. Indirectness: No indirectness

Funding	Equipment / drugs provided by industry (Thyrogen medication was provided by Genzyme Corp. Other than Tyrogen medication there was no financial support or other support whatsoever by internal, external, government or industry.)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: RHTSH versus Withdrawal Protocol outcome 1: Cancer recurrence - Actual outcome for Ablation: cancer recurrence up to 4.5 years follow up. Group 1: 1/24, Group 2: 0/20 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: 0. Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Quality of life . Mortality . Local cancer progression . Incidence of distant metastases . Successful ablation

Study	ESTIMABL1 trial: Schlumberger 2012[37], 4, 2015[7]
Study type	RCT (Patient randomised. Parallel)
Number of studies (number of participants)	(n=752)
Countries and setting	Conducted in France. Setting: 24 French medical centers
Line of therapy	Not applicable
Duration of study	Intervention + follow up: 5 years
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Ablation
Subgroup analysis within study	Not applicable
Inclusion criteria	an age of 18 years or older, low-risk differentiated thyroid carcinoma (papillary or follicular, excluding aggressive histologic subtypes),25 tumor–node–metastasis (TNM) stage, ascertained on pathological examination (p) of a surgical specimen, of pT1 (tumor diameter ≤1 cm) and N1 or Nx or pT1 (tumor diameter >1 to 2 cm) and any N or pT2N0,26 absence of distant metastasis, Eastern Cooperative Oncology Group performance status score of 0 or 1 (i.e., fully active and able to carry on all predisease performance without restriction, and restricted from physically strenuous activity but ambulatory, respectively), no major coexisting conditions (including other cancers) within the previous 5 years, and a negative pregnancy test for women.
Exclusion criteria	Patients with a recent history of iodine contamination were excluded.
Recruitment/selection of patients	patients with low-risk thyroid cancer after a complete surgical resection.
Age, gender and ethnicity	Age - Mean (SD): rhTSH: 1.1GBq 51±13. 3.7GBq 48±14. Withdrawal: 1.1GBq 49±13. 3.7GBq 49±14. Gender (M:F): 162/590. Ethnicity:
Further population details	1. Activity level: Activity higher (3 - 4 Gb) (1.1GBq: 376. 3.3GBq: 376). 2. Diet: Not stated / Unclear 3. TSH levels: Not stated / Unclear
Extra comments	The study compared two thyrotropin-stimulation methods (thyroid hormone withdrawal and use of recombinant human thyrotropin) and two radioiodine (131I) doses (1.1 GBq and 3.7 GBq) in a 2-by-2 design. For the purposes of this review the two thyrotropin stimulation methods were compared.
Indirectness of population	--
Interventions	(n=374) Intervention 1: Radioactive iodine ablation - with prior preparation with thyrotropin alfa. All patients underwent total thyroidectomy. Lymph-node dissection was performed in patients with evidence of lymph-node involvement, as well as in some patients with no evidence of lymph node involvement, if part of local practice. Randomization was performed between 30 and 120 days after surgery, during which time patients received

	levothyroxine therapy for at least 28 days (or levotri-iodothyronine therapy for 14 days). Recombinant human thyrotropin (Thyrogen, Genzyme) was administered during treatment with thyroid hormone, at a dose of 0.9 mg intra-muscularly on 2 consecutive days, and radioiodine was administered on the day after the second injection.. Duration 28 days. Concurrent medication/care: one of two 131I activities (1.1 GBq or 3.7 GBq). . Indirectness: No indirectness (n=378) Intervention 2: Radioactive iodine ablation - with prior withdrawal of thyroid hormone replacement. All patients underwent total thyroidectomy. Lymph-node dissection was performed in patients with evidence of lymph-node involvement, as well as in some patients with no evidence of lymph node involvement, if part of local practice. Thyroid-hormone withdrawal consisted of discontinuation of levothyroxine treatment for at least 28 days (or levotriiodothyronine treatment withdrawal for 14 days), with administration of radioiodine when the serum thyrotropin concentration was higher than 30 mIU per liter. . Duration 14 - 28 days. Concurrent medication/care: one of two 131I activities (1.1 GBq or 3.7 GBq). . Indirectness: No indirectness
Funding	Academic or government funding (Funded by the French National Cancer Institute [INCa] and the French Ministry of Health. ClinicalTrials.gov number, NCT00435851. INCa number, RECF0447)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: RHTSH versus WITHDRAWAL

Protocol outcome 1: Successful ablation
- Actual outcome for Ablation: Thyroglobulin ≤1ng/ml at 6-10 months post RAI. Group 1: 317/334, Group 2: 304/318
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: 14, Reason: specific breakdown not given. patients lost to follow up, ongoing disease and did not undergo treatment. Group 2 Number missing: 18, Reason: specific breakdown not given. patients lost to follow up, ongoing disease and did not undergo treatment
- Actual outcome for Ablation: Complete ablation at 6-10 months post RAI. Group 1: 319/348, Group 2: 312/336. Comments: Ablation was considered complete if both the neck ultrasound was normal and the level of recombinant human thyrotropin-stimulated thyroglobulin was less than or equal to 1ng/ml (or in cases of detectable antithyroglobulin antibody if the control 131I total body scan was normal.
Risk of bias: All domain – Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 0, Reason: no of participants included in analysis not provided. Group 2 Number missing: 0, Reason: no of participants included in analysis not provided

Protocol outcome 1: Quality of life
- Actual outcome for quality of life: EQ5D utility score mean (sd). Group 1: 0.849 (0.173), Group 2: 0.833 (0.192)
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: 14, Reason: specific breakdown not given. patients lost to follow up, ongoing

disease and did not undergo treatment. Group 2 Number missing: 18, Reason: specific breakdown not given. patients lost to follow up, ongoing disease and did not undergo treatment

- Actual outcome for quality of life: SF36 physical functioning at 1.1GBq . Group 1: 86 (17), Group 2: 79 (20). Comments: Data stratified for RAI activity level, and there was no summated analysis.
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 0, Reason: no of participants included in analysis not provided. Group 2 Number missing: 0, Reason: no of participants included in analysis not provided

- Actual outcome for quality of life: SF36 role physical at 1.1GBq . Group 1: 75 (26), Group 2: 61 (30). Comments: Data stratified for RAI activity level, and there was no summated analysis.
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 0, Reason: no of participants included in analysis not provided. Group 2 Number missing: 0, Reason: no of participants included in analysis not provided

- Actual outcome for quality of life: SF36 bodily pain at 1.1GBq . Group 1: 77 (23), Group 2: 70 (25). Comments: Data stratified for RAI activity level, and there was no summated analysis.
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 0, Reason: no of participants included in analysis not provided. Group 2 Number missing: 0, Reason: no of participants included in analysis not provided

- Actual outcome for quality of life: SF36 general health at 1.1GBq . Group 1: 67 (17), Group 2: 65 (19). Comments: Data stratified for RAI activity level, and there was no summated analysis.
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 0, Reason: no of participants included in analysis not provided. Group 2 Number missing: 0, Reason: no of participants included in analysis not provided

- Actual outcome for quality of life: SF36 vitality at 1.1GBq . Group 1: 54 (22), Group 2: 43 (24). Comments: Data stratified for RAI activity level, and there was no summated analysis.
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 0, Reason: no of participants included in analysis not provided. Group 2 Number missing: 0, Reason: no of participants included in analysis not provided

- Actual outcome for quality of life: SF36 social functioning at 1.1GBq . Group 1: 76 (24), Group 2: 65 (28). Comments: Data stratified for RAI activity level, and there was no summated analysis.
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 0, Reason: no of participants included in analysis not provided. Group 2 Number missing: 0, Reason: no of participants included in analysis not provided

- Actual outcome for quality of life: SF36 emotional role at 1.1GBq . Group 1: 78 (24), Group 2: 70 (23). Comments: Data stratified for RAI activity level, and there was no summated analysis.
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 0, Reason: no of participants included in analysis not provided. Group 2 Number missing: 0, Reason: no of participants included in analysis not provided

- Actual outcome for quality of life: SF36 mental health at 1.1GBq . Group 1: 66 (21), Group 2: 65 (20). Comments: Data stratified for RAI activity level, and there was no summated analysis.
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 0, Reason: no of participants included in analysis not provided. Group 2 Number missing: 0, Reason: no of participants included in analysis not provided

- Actual outcome for quality of life: SF36 mental summary component at 1.1GBq . Group 1: 44 (12), Group 2: 41 (12). Comments: Data stratified for RAI activity level, and there was no summated analysis.
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 0, Reason: no of participants included in analysis not provided. Group 2 Number missing: 0, Reason: no of participants included in analysis not provided

- Actual outcome for quality of life: SF36 physical summary component at 1.1GBq . Group 1: 52 (7), Group 2: 48 (9). Comments: Data stratified for RAI activity level, and there was no summated analysis.
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 0, Reason: no of participants included in analysis not provided. Group 2 Number missing: 0, Reason: no of participants included in analysis not provided

- Actual outcome for quality of life: SF36 physical functioning at 3.7GBq . Group 1: 86 (17), Group 2: 78 (22). Comments: Data stratified for RAI activity level, and there was no summated analysis.
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 0, Reason: no of participants included in analysis not provided. Group 2 Number missing: 0, Reason: no of participants included in analysis not provided

- Actual outcome for quality of life: SF36 role physical at 3.7GBq . Group 1: 75 (27), Group 2: 59 (29). Comments: Data stratified for RAI activity level, and there was no summated analysis.
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 0, Reason: no of participants included in analysis not provided. Group 2 Number missing: 0, Reason: no of participants included in analysis not provided

- Actual outcome for quality of life: SF36 bodily pain at 3.7GBq . Group 1: 77 (23), Group 2: 69 (27). Comments: Data stratified for RAI activity level, and

there was no summated analysis.
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 0, Reason: no of participants included in analysis not provided. Group 2 Number missing: 0, Reason: no of participants included in analysis not provided

- Actual outcome for quality of life: SF36 general health at 3.7GBq . Group 1: 66 (17), Group 2: 64 (20). Comments: Data stratified for RAI activity level, and there was no summated analysis.
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 0, Reason: no of participants included in analysis not provided. Group 2 Number missing: 0, Reason: no of participants included in analysis not provided

- Actual outcome for quality of life: SF36 vitality at 3.7GBq . Group 1: 56 (22), Group 2: 42 (23). Comments: Data stratified for RAI activity level, and there was no summated analysis.
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 0, Reason: no of participants included in analysis not provided. Group 2 Number missing: 0, Reason: no of participants included in analysis not provided

- Actual outcome for quality of life: SF36 social functioning at 3.7GBq . Group 1: 78 (23), Group 2: 66 (26). Comments: Data stratified for RAI activity level, and there was no summated analysis.
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 0, Reason: no of participants included in analysis not provided. Group 2 Number missing: 0, Reason: no of participants included in analysis not provided

- Actual outcome for quality of life: SF36 emotional role at 3.7GBq . Group 1: 78 (26), Group 2: 70 (27). Comments: Data stratified for RAI activity level, and there was no summated analysis.
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 0, Reason: no of participants included in analysis not provided. Group 2 Number missing: 0, Reason: no of participants included in analysis not provided

- Actual outcome for quality of life: SF36 mental health at 3.7GBq . Group 1: 66 (21), Group 2: 64 (22). Comments: Data stratified for RAI activity level, and there was no summated analysis.
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 0, Reason: no of participants included in analysis not provided. Group 2 Number missing: 0, Reason: no of participants included in analysis not provided

- Actual outcome for quality of life: SF36 mental summary component at 3.7GBq . Group 1: 44 (12), Group 2: 41 (13). Comments: Data stratified for RAI activity level, and there was no summated analysis.
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover -

Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 0, Reason: no of participants included in analysis not provided. Group 2 Number missing: 0, Reason: no of participants included in analysis not provided - Actual outcome for quality of life: SF36 physical summary component at 3.7GBq . Group 1: 52 (8), Group 2: 47 (10). Comments: Data stratified for RAI activity level, and there was no summated analysis. Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 0, Reason: no of participants included in analysis not provided. Group 2 Number missing: 0, Reason: no of participants included in analysis not provided	
Protocol outcomes not reported by the study	Mortality . Local cancer progression . Incidence of distant metastases . Cancer recurrence

Study (subsidiary papers)	HiLo Trial trial: Mallick 2012[25] merged with Dehbi 2019[12]
Study type	RCT (Patient randomised. Parallel)
Number of studies (number of participants)	1 (n=438)
Countries and setting	Conducted in United Kingdom. Setting:
Line of therapy	Not applicable
Duration of study	Intervention + follow up: 6-9 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Ablation
Subgroup analysis within study	Not applicable
Inclusion criteria	Eligibility criteria were an age of 16 to 80 years, a performance status of 0 to 2 (with 0 indicating normal function, 1 indicating that the patient is restricted in strenuous activity but ambulatory, and 2 indicating that the patient is capable of self-care but is unable to work), histological confirmation of differentiated thyroid cancer (including Hürthle-cell carcinoma) requiring radioiodine ablation. tumor stage T1 to T3 with the possibility of lymph-node involvement but no distant metastasis and no microscopical residual disease (i.e., N0, NX, N1, and M0 in the tumor–node–metastasis [TNM sixth] staging system), and one- or two-stage total thyroidectomy, with or without central lymph-node dissection.
Exclusion criteria	Exclusion criteria were the presence of aggressive malignant variants, including tall-cell, insular, poorly differentiated, and diffuse sclerosing thyroid cancer. anaplastic or medullary carcinoma. pregnancy. severe coexisting conditions. previous cancer with limited life expectancy. previous iodine-131 or iodine-123 preablation scanning. and previous treatment for thyroid cancer except surgery.
Recruitment/selection of patients	Patients requiring radioiodine ablation after total thyroidectomy
Age, gender and ethnicity	Age - Median (range): rhTSH: 44 (20-82) / 44 (21-76). Withdrawal: 45 (17-73) / 43 (18-77). Gender (M:F): 111/326. Ethnicity:
Further population details	1. Activity level: Activity higher (3 - 4 Gb) (3.7GBq - 220. 3.7 GBq - 218). 2. Diet: Additional dietary restrictions (All patients were instructed to follow a low iodine diet for 3 weeks before ablation). 3. TSH levels: Not stated / Unclear
Extra comments	In this study, patients were randomly assigned to one of four study groups: low-dose or high-dose radioiodine, each combined with thyrotropin alfa (Thyrogen, Genzyme) or thyroid hormone withdrawal. For the purposes of this review, outcomes which have combined the low and high dose radioiodine and compared rhTSH to withdrawal have been used.

Indirectness of population	No indirectness
Interventions	(n=219) Intervention 1: Radioactive iodine ablation - with prior preparation with thyrotropin alfa. Thyrotropin alfa was administered on each of the 2 days before ablation by intramuscular injection (0.9 mg).. Duration 2 days prior to ablation. Concurrent medication/care: Radioactive iodine-131 was administered at a dose of 1.1 GBq (n=110) or 3.7 GBq (n=109), depending on the study group.. Indirectness: No indirectness (n=219) Intervention 2: Radioactive iodine ablation - with prior withdrawal of thyroid hormone replacement. Among the patients undergoing thyroid hormone withdrawal, thyroxine (average dose, 200 µg per day) was discontinued 4 weeks before ablation in 11 patients, and triiodothyronine (average dose, 60 µg per day) was discontinued for 2 weeks in 204 patients.. Duration 2 - 4 weeks prior to ablation. Concurrent medication/care: Radioactive iodine-131 was administered at a dose of 1.1 GBq or 3.7 GBq, depending on the study group.. Indirectness: No indirectness
Funding	Academic or government funding (Supported by grants from Cancer Research UK (C18243/A5802) and University College London and the University College London Hospital Comprehensive Biomedical Research Centre.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: RHTSH versus WITHDRAWAL

Protocol outcome 1: Quality of life
- Actual outcome for Ablation: SF-36 (psychological domains) at 3 months post ablation. Group 1: mean 24 (SD 109.98). n=219, Group 2: mean 19 (SD 109.98). n=219. SF-36 0-100 Top=High is good outcome
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 0, Reason: Patients were excluded from each comparison if they had neither diagnostic scanning nor thyroglobulin testing.. Group 2 Number missing: 0, Reason: Patients were excluded from each comparison if they had neither diagnostic scanning nor thyroglobulin testing.
- Actual outcome for Ablation: SF-36 (physical domains) at 3 months post ablation. Group 1: mean 15.6 (SD 108.38). n=219, Group 2: mean 17.5 (SD 108.38). n=219. SF-36 0-100 Top=High is good outcome
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 0, Reason: Patients were excluded from each comparison if they had neither diagnostic scanning nor thyroglobulin testing.. Group 2 Number missing: 0, Reason: Patients were excluded from each comparison if they had neither diagnostic scanning nor thyroglobulin testing.
- Actual outcome for Ablation: SF-36 (general health) at 3 months post ablation. Group 1: mean -0.6 (SD 24.02). n=219, Group 2: mean -1.7 (SD 24.025). n=219. SF-36 0-100 Top=High is good outcome
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 0, Reason: Patients were excluded from each comparison if they had neither diagnostic scanning nor thyroglobulin testing.. Group 2 Number missing: 0, Reason: Patients were excluded from each comparison if they had neither

diagnostic scanning nor thyroglobulin testing.
- Actual outcome for Ablation: SF-36 (physical functioning) at 3 months post ablation. Group 1: mean 0.5 (SD 31.5). n=219, Group 2: mean -0.6 (SD 31.5). n=219. SF-36 0-100 Top=High is good outcome
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 0, Reason: Patients were excluded from each comparison if they had neither diagnostic scanning nor thyroglobulin testing.. Group 2 Number missing: 0, Reason: Patients were excluded from each comparison if they had neither diagnostic scanning nor thyroglobulin testing.
- Actual outcome for Ablation: SF-36 (role limitations due to physical functioning) at 3 months post ablation. Group 1: mean 10 (SD 44.84). n=219, Group 2: mean 15.1 (SD 44.84). n=219. SF-36 0-100 Top=High is good outcome
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 0, Reason: Patients were excluded from each comparison if they had neither diagnostic scanning nor thyroglobulin testing.. Group 2 Number missing: 0, Reason: Patients were excluded from each comparison if they had neither diagnostic scanning nor thyroglobulin testing.
- Actual outcome for Ablation: SF-36 (role limitations due to emotional problems) at 3 months post ablation. Group 1: mean 5.4 (SD 40.04). n=219, Group 2: mean 2.2 (SD 40.04). n=219. SF-36 0-100 Top=High is good outcome
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 0, Reason: Patients were excluded from each comparison if they had neither diagnostic scanning nor thyroglobulin testing.. Group 2 Number missing: 0, Reason: Patients were excluded from each comparison if they had neither diagnostic scanning nor thyroglobulin testing.
- Actual outcome for Ablation: SF-36 (social functioning) at 3 months post ablation. Group 1: mean 7.7 (SD 38.44). n=219, Group 2: mean 8.8 (SD 38.44). n=219. SF-36 0-100 Top=High is good outcome
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 0, Reason: Patients were excluded from each comparison if they had neither diagnostic scanning nor thyroglobulin testing.. Group 2 Number missing: 0, Reason: Patients were excluded from each comparison if they had neither diagnostic scanning nor thyroglobulin testing.
- Actual outcome for Ablation: SF-36 (Pain) at 3 months post ablation. Group 1: mean 5.4 (SD 40.04). n=219, Group 2: mean 5.5 (SD 40.04). n=219. SF-36 0-100 Top=High is good outcome
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 0, Reason: Patients were excluded from each comparison if they had neither diagnostic scanning nor thyroglobulin testing.. Group 2 Number missing: 0, Reason: Patients were excluded from each comparison if they had neither diagnostic scanning nor thyroglobulin testing.
- Actual outcome for Ablation: SF-36 (Energy / Fatigue) at 3 months post ablation. Group 1: mean 4.5 (SD 32.03). n=219, Group 2: mean 4.1 (SD 32.09). n=219. SF-36 0-100 Top=High is good outcome
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 0, Reason: Patients were excluded from each comparison if they had neither diagnostic scanning nor thyroglobulin testing.. Group 2 Number missing: 0, Reason: Patients were excluded from each comparison if they had neither diagnostic scanning nor thyroglobulin testing.
- Actual outcome for Ablation: SF-36 (Emotional well-being) at 3 months post ablation. Group 1: mean 4.1 (SD 27.22). n=219, Group 2: mean 3 (SD 27.22).

n=219. SF-36 0-100 Top=High is good outcome Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness. Group 1 Number missing: 0, Reason: Patients were excluded from each comparison if they had neither diagnostic scanning nor thyroglobulin testing.. Group 2 Number missing: 0, Reason: Patients were excluded from each comparison if they had neither diagnostic scanning nor thyroglobulin testing. Protocol outcome 2: Cancer recurrence - Actual outcome for Ablation: Cancer recurrence at median follow-up was 6·5 years (IQR 4·5–7·6). Group 1: 13/218, Group 2: 8/216 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: 1, Reason: lost to follow up. Group 2 Number missing: 3, Reason: lost to follow up Protocol outcome 3: Successful ablation - Actual outcome for Ablation: Ablation success based on diagnostic scan alone (<0.1%) at 6 - 9 months post treatment. Group 1: 197/210, Group 2: 198/211 Risk of bias: All domain - Low, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 9, Reason: Patients were excluded from each comparison if they had neither diagnostic scanning nor thyroglobulin testing.. Group 2 Number missing: 8, Reason: Patients were excluded from each comparison if they had neither diagnostic scanning nor thyroglobulin testing. - Actual outcome for Ablation: Ablation success based on Thyroglobulin levels alone (<0.2ng/nl) at 6 - 9 months post treatment. Group 1: 162/185, Group 2: 150/174 Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 34, Reason: Patients were excluded from each comparison if they had neither diagnostic scanning nor thyroglobulin testing.. Group 2 Number missing: 45, Reason: Patients were excluded from each comparison if they had neither diagnostic scanning nor thyroglobulin testing. - Actual outcome for Ablation: Ablation success based on Thyroglobulin levels (<0.2ng/nl) and diagnostic scan (<0.1%) at 6 - 9 months post treatment. Group 1: 183/210, Group 2: 183/211 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: 9, Reason: Patients were excluded from each comparison if they had neither diagnostic scanning nor thyroglobulin testing.. Group 2 Number missing: 8, Reason: Patients were excluded from each comparison if they had neither diagnostic scanning nor thyroglobulin testing.	
Protocol outcomes not reported by the study	Mortality . Local cancer progression . Incidence of distant metastases

Study	Lee 2010[21]
Study type	RCT (Patient randomised. Parallel)
Number of studies (number of participants)	(n=291)
Countries and setting	Conducted in South Korea. Setting: Medical university hospitals in the Republic of Korea
Line of therapy	Not applicable
Duration of study	Intervention + follow up: 12 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Ablation
Subgroup analysis within study	Not applicable
Inclusion criteria	Patients with newly diagnosed disseminated thyroid cancer, more than 18 years old, who had recently undergone total or near total thyroidectomy with central compartment neck dissection.
Exclusion criteria	Evidence of distant metastases, lateral neck node metastases, and or significant extra thyroidal invasion. Included patients had no clinically significant abnormalities on routine haematological or blood chemistry tests, and serum creatinine concentrations were normal. No patient had any major concurrent medical disorders, including other malignancies, within the past 5 years, and no patient had recently been prescribed drugs affecting thyroid or renal function, including iodine containing medications or radiocontrast agents.
Recruitment/selection of patients	Patients undergoing radioiodine ablation treatment
Age, gender and ethnicity	Age - Mean (SD): rhTSH: 46.7 ± 9.8. Withdrawal: 50.1 ± 6.8. Gender (M:F): 11/147. Ethnicity:
Further population details	1. Activity level: Activity low (1Gb) (30mCi or 1.11 GBq). 2. Diet: Additional dietary restrictions (low iodine diet for two weeks prior to treatment). 3. TSH levels: High (>30) (rhTSH: 86.6 ± 17.6 mU/L. withdrawal: 81.2 ± 19 mU/L).
Extra comments	The study has a third comparison arm consisting of patients who discontinued levothyroxine for 4 weeks plus 2 weeks on and then 2 weeks off liothyronine. This data has not been extracted for the purposes of this review.
Indirectness of population	No indirectness
Interventions	(n=69) Intervention 1: Radioactive iodine ablation - with prior preparation with thyrotropin alfa. All patients underwent total thyroidectomy with central compartment neck dissection. After the operation, all patients began treatment with TSH supressing dose of LT4 (levothyroxine 2µg / kg) after at least 30 days of LT4 supplementation, patients were randomized into groups. in the rhTSH group, each patient received two injections of rhTSH: 0.9mg IM at 24 hours and 48 hours before the administration of the RI therapeutic dose. Duration 48 hours prior to radioiodine ablation. Concurrent medication/care: remnant ablation using low dose

	(30 mCi / 1.11GBq) radioiodine treatment. Indirectness: No indirectness (n=89) Intervention 2: Radioactive iodine ablation - with prior withdrawal of thyroid hormone replacement. All patients underwent total thyroidectomy with central compartment neck dissection. After the operation, all patients began treatment with TSH supressing dose of LT4 (levothyroxine 2μg / kg) after at least 30 days of LT4 supplementation, patients were randomized into groups. Those in the T4 withdrawal group discontinued LT4 for 4 weeks.. Duration 4 weeks prior to radioiodine treatment. Concurrent medication/care: remnant ablation using low dose (30 mCi / 1.11GBq) radioiodine treatment. Indirectness: No indirectness
Funding	Academic or government funding (the study was supported by research funds of Yonsei University College of Medicine in 20016)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: RHTSH versus WITHDRAWAL

Protocol outcome 1: Incidence of distant metastases
- Actual outcome for Ablation: Lymph node metastases at 12 months after radioiodine treatment. Group 1: 3/69, Group 2: 2/89
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: 0. Group 2 Number missing: 0

Protocol outcome 2: Successful ablation
- Actual outcome for Ablation: Successful ablation (no visible uptake or below 0.1%) at 12 months after radioiodine treatment. Group 1: 63/69, Group 2: 83/89
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: 0. Group 2 Number missing: 0
- Actual outcome for Ablation: serum thyroglobulin ≤1.0ng/mL at 12 months after radioiodine treatment. Group 1: 64/69, Group 2: 81/89
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: 0. Group 2 Number missing: 0
- Actual outcome for Ablation: Thyroglobulin levels at 12 months after radioiodine treatment. Group 1: mean 0.14 ng/mL (SD 0.05). n=69, Group 2: mean 0.18 ng/mL (SD 0.14). n=89
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: 0. Group 2 Number missing: 0

Protocol outcomes not reported by the study	Quality of life . Mortality . Local cancer progression . Cancer recurrence

Study (subsidiary papers)	Pacini 2006[33] merged with Hanscheid 2006[16]
Study type	RCT (Patient randomised. Parallel)
Number of studies (number of participants)	(n=63)
Countries and setting	Conducted in Multiple countries. Setting: Four centers in Europe and five in North America.
Line of therapy	Not applicable
Duration of study	Intervention + follow up: 8 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Ablation
Subgroup analysis within study	Not applicable
Inclusion criteria	Study patients were 18 yr or older with newly diagnosed differentiated papillary or follicular thyroid carcinoma, the sole previous treatment for which had been total or near-total thyroidectomy within 2 wk before enrolment. Patients had no clinically significant abnormalities of haematological or blood chemistry testing for routine analytes, including serum creatinine concentration. No patients had major concurrent medical disorders, including other malignancies within the past 5 yr. and no patient had a recent history of drugs affecting thyroid or renal function, including iodine-containing medications or radiocontrast agents.
Exclusion criteria	not specified
Recruitment/selection of patients	Patients were all staged T2 or T4 with minor invasion of the thyroid capsule, N0-N1, and M0 or T0-T1, N1, and M0. T4 tumors were no longer eligible after a protocol amendment because concern arose that patients with T4 tumors might alternatively be treated routinely with radioiodine doses higher than 100 mCi or external radiotherapy at some centers. However, six T4 patients already enrolled before the study amendment
Age, gender and ethnicity	Age - Mean (SD): Withdrawal: 43.2 (12.5). rhTSH: 44.5 (12.2). Gender (M:F): 13/50. Ethnicity:
Further population details	1. Activity level: Activity higher (3 - 4 Gb) (3.7GBq(100 mCi)). 2. Diet: No additional dietary restrictions 3. TSH levels: High (>30) (rhTSH: 1.1 ± 1.3 mU/liter. hypothyroid: 83 ± 51 mU/liter).
Indirectness of population	No indirectness
Interventions	(n=33) Intervention 1: Radioactive iodine ablation - with prior preparation with thyrotropin alfa. Patients in the euthyroid group received l-thyroxine therapy for 4–6 wk until their serum TSH concentration was 5 mU/liter or less. Then 0.9 mg rhTSH (Thyrogen) was administered im on 2 consecutive days.. Duration 4 - 6 weeks prior to administration. Concurrent medication/care: 24 h after rhTSH, 3.7 GBq (100 mCi) 131I was administered.. Indirectness: No indirectness

	(n=30) Intervention 2: Radioactive iodine ablation - with prior withdrawal of thyroid hormone replacement. Patients randomized to the hypothyroid group did not receive thyroid hormone therapy postoperatively. The serum TSH concentration was reassessed at 4–6 wk until the patient's TSH was greater than 25mU/liter.. Duration 4 - 6 weeks prior to RAI treatment. Concurrent medication/care: The patients received a 3.7GBq(100 mCi) 131 . Indirectness: No indirectness
Funding	Other (This work was supported by the Genzyme Corp. (Cambridge, MA).)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: RHTSH versus Withdrawal

Protocol outcome 1: Quality of life
- Actual outcome for Ablation: SF 36 (mental component) at 4 weeks post treatment. Group 1: mean 45.2 (SD 11.9). n=33, Group 2: mean 38.5 (SD 9.8). n=30. SF-36 0-100 Top=High is good outcome. Comments:
Baseline results:
Hypothyroid: 44.4 ± 12.0
rhTSH: 40 ± 10.0
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: 1, Reason: one patient was ineligible for the final analysis due to a mistake in the reconstitution of one rhTSH dose in preparation for 131I ablation. Group 2 Number missing: 2, Reason: one due to discovery of lung metastases on the post therapy whole-body scan and one because the neck scan was uninterpretable due to a positioning error
- Actual outcome for Ablation: SF 36 (physical component) at 4 weeks post treatment. Group 1: mean 47.6 (SD 7.7). n=33, Group 2: mean 40 (SD 9.9). n=30. SF-36 0-100 Top=High is good outcome. Comments:
Baseline results:
rhTSH: 46.2 ± 7.5
Hypothyroid: 42.5 ± 7.2
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: 1, Reason: one patient was ineligible for the final analysis due to a mistake in the reconstitution of one rhTSH dose in preparation for 131I ablation. Group 2 Number missing: 2, Reason: one due to discovery of lung metastases on the post therapy whole-body scan and one because the neck scan was uninterpretable due to a positioning error
- Actual outcome for Ablation: SF 36 (physical functioning) at 4 weeks post treatment. Group 1: mean 84.5 (SD 18.3). n=33, Group 2: mean 57.8 (SD 29.4). n=30. SF-36 0-100 Top=High is good outcome. Comments:
Baseline results:
rhTSH: 82.0 ± 18.5
Hypothyroid: 71.0 ± 26.5
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: 1, Reason: one patient was ineligible for the final analysis due to a mistake in the reconstitution of one rhTSH dose in preparation for 131I ablation. Group 2 Number missing: 2, Reason: one due to discovery of lung metastases on the post

therapy whole-body scan and one because the neck scan was uninterpretable due to a positioning error
- Actual outcome for Ablation: SF 36 (role - physical) at 4 weeks post treatment. Group 1: mean 58.3 (SD 38.9). n=33, Group 2: mean 22.5 (SD 34.3). n=30. SF-36 0-100 Top=High is good outcome. Comments:
Baseline results:
rhTSH: 43 ± 44.6
Hypothyroid: 36.7 ± 36.4
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: 1, Reason: one patient was ineligible for the final analysis due to a mistake in the reconstitution of one rhTSH dose in preparation for 131I ablation. Group 2 Number missing: 2, Reason: one due to discovery of lung metastases on the post therapy whole-body scan and one because the neck scan was uninterpretable due to a positioning error
- Actual outcome for Ablation: SF 36 (bodily pain) at 4 weeks post treatment. Group 1: mean 67.4 (SD 23.6). n=33, Group 2: mean 55 (SD 22.4). n=30. SF-36 0-100 Top=High is good outcome. Comments:
Baseline results:
rhTSH: 57.8 ± 28.3
Hypothyroid: 54.1 ± 27.1Risk 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: 1, Reason: one patient was ineligible for the final analysis due to a mistake in the reconstitution of one rhTSH dose in preparation for 131I ablation. Group 2 Number missing: 2, Reason: one due to discovery of lung metastases on the post therapy whole-body scan and one because the neck scan was uninterpretable due to a positioning error
- Actual outcome for Ablation: SF 36 (general health) at 4 weeks post treatment. Group 1: mean 66.1 (SD 20.8). n=33, Group 2: mean 61.6 (SD 21.2). n=30. SF-36 0-100 Top=High is good outcome. Comments:
Baseline results:
rhTSH: 68.2 ± 18.4
Hypothyroid: 67.8 ± 15.1
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: 1, Reason: one patient was ineligible for the final analysis due to a mistake in the reconstitution of one rhTSH dose in preparation for 131I ablation. Group 2 Number missing: 2, Reason: one due to discovery of lung metastases on the post therapy whole-body scan and one because the neck scan was uninterpretable due to a positioning error
- Actual outcome for Ablation: SF 36 (Vitality) at 4 weeks post treatment. Group 1: mean 54.5 (SD 22.5). n=33, Group 2: mean 36.4 (SD 21.3). n=30. SF-36 0-100 Top=High is good outcome. Comments:
Baseline results:
rhTSH: 46.6 ± 22.2
Hypothyroid:55.7 ± 23.3
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: 1, Reason: one patient was ineligible for the final analysis due to a mistake in the reconstitution of one rhTSH dose in preparation for 131I ablation. Group 2 Number missing: 2, Reason: one due to discovery of lung metastases on the post therapy whole-body scan and one because the neck scan was uninterpretable due to a positioning error
- Actual outcome for Ablation: SF 36 (social functioning) at 4 weeks post treatment. Group 1: mean 74.2 (SD 21.4). n=33, Group 2: mean 53.3 (SD 28.4). n=30. SF-36 0-100 Top=High is good outcome. Comments:

Baseline results:
rhTSH: 62.1 ± 24.3
Hypothyroid: 67.5 ± 24.5
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: 1, Reason: one patient was ineligible for the final analysis due to a mistake in the reconstitution of one rhTSH dose in preparation for 131I ablation. Group 2 Number missing: 2, Reason: one due to discovery of lung metastases on the post therapy whole-body scan and one because the neck scan was uninterpretable due to a positioning error
- Actual outcome for Ablation: SF 36 (role - emotional) at 4 weeks post treatment. Group 1: mean 57.6 (SD 44.3). n=33, Group 2: mean 31.1 (SD 41). n=30. SF-36 0-100 Top=High is good outcome. Comments:
Baseline results:
rhTSH: 46.9 ± 43.9
Hypothyroid: 50 ± 44.4
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: 1, Reason: one patient was ineligible for the final analysis due to a mistake in the reconstitution of one rhTSH dose in preparation for 131I ablation. Group 2 Number missing: 2, Reason: one due to discovery of lung metastases on the post therapy whole-body scan and one because the neck scan was uninterpretable due to a positioning error
- Actual outcome for Ablation: SF 36 (mental health) at 4 weeks post treatment. Group 1: mean 71 (SD 20.1). n=33, Group 2: mean 58.8 (SD 16.5). n=33. SF-36 0-100 Top=High is good outcome. Comments:
Baseline results:
rhTSH: 61.4 ± 18.8
Hypothyroid: 64.3 ± 18.4
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: 1, Reason: one patient was ineligible for the final analysis due to a mistake in the reconstitution of one rhTSH dose in preparation for 131I ablation. Group 2 Number missing: 2, Reason: one due to discovery of lung metastases on the post therapy whole-body scan and one because the neck scan was uninterpretable due to a positioning error

Protocol outcome 2: Successful ablation
- Actual outcome for Ablation: No visible uptake at 8 months post treatment. Group 1: 24/32, Group 2: 24/28
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: 1, Reason: one patient was ineligible for the final analysis due to a mistake in the reconstitution of one rhTSH dose in preparation for 131I ablation. Group 2 Number missing: 2, Reason: one due to discovery of lung metastases on the post therapy whole-body scan and one because the neck scan was uninterpretable due to a positioning error
- Actual outcome for Ablation: Visible uptake <0.1% at 8 months post treatment. Group 1: 4/28, Group 2: 8/32
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - High, Measurement - Low, Crossover - Low. Indirectness of outcome: No indirectness . Group 1 Number missing: 1, Reason: one patient was ineligible for the final analysis due to a mistake in the reconstitution of one rhTSH dose in preparation for 131I ablation. Group 2 Number missing: 2, Reason: one due to discovery of lung metastases on the post therapy whole-body scan and one because the neck scan was uninterpretable due to a positioning error

Protocol outcomes not reported by the study	Mortality . Local cancer progression . Incidence of distant metastases . Cancer recurrence

Study	Taieb 2009[39]
Study type	RCT (Patient randomised. Parallel)
Number of studies (number of participants)	(n=74)
Countries and setting	Conducted in France. Setting: not specified
Line of therapy	Not applicable
Duration of study	Intervention + follow up: 9 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Ablation
Subgroup analysis within study	Not applicable
Inclusion criteria	Aged ≥18 years, newly diagnosed well differentiated papillary or follicular carcinoma in patients who had total thyroidectomy (one stage or two stage), all staged pT1-T3, N0-Nx-N1, M0 (if <5 nodes and without extracapsular spread) and all patients gave their signed consent.
Exclusion criteria	presence of distant metastases, previous history of major concurrent chronic medical disorders, psychiatric disorders, chronic alcoholism and external radiotherapy or malignancies
Recruitment/selection of patients	Inclusion of patients was performed the day after thyroidectomy (total thyroidectomy or two stage completion thyroidectomy) by endocrine surgeons
Age, gender and ethnicity	Age - Mean (SD): Withdrawal: 49 ± 11.8. rhTSH: 45.5 ± 15.6. Gender (M:F): 12/62. Ethnicity:
Further population details	1. Activity level: Activity higher (3 - 4 Gb) (3.7GBq). 2. Diet: Not stated / Unclear 3. TSH levels: Not stated / Unclear
Indirectness of population	No indirectness
Interventions	(n=37) Intervention 1: Radioactive iodine ablation - with prior withdrawal of thyroid hormone replacement. Patients were discharged from the department of endocrine surgery with levothyroxine supplementation (2µg/kg). One week later patients were randomized into the hypo group in which patients discontinued L-T4 for 5 weeks. Duration 5 - 6 weeks. Concurrent medication/care: All patients received 3.7GBq activity at 6 weeks post surgery. Indirectness: No indirectness (n=37) Intervention 2: Radioactive iodine ablation - with prior preparation with thyrotropin alfa. Patients were discharged from the department of endocrine surgery with levothyroxine supplementation (2µg/kg). One week later patients were randomized into the rhTSH group in which patients continued to take L-T4 and received rhTSH (two 0.9mg IM injections on two consecutive days as ambulatory patients) 1 - 2 weeks later. Both injections were performed at the institution to ensure injection and TSH peak was validated. . Duration 2 - 3

	weeks. Concurrent medication/care: All patients received 3.7GBq activity at 2 - 3 weeks post surgery. Indirectness: No indirectness
Funding	Academic or government funding (This work was financially supported by the Genzyme Corporation (Cambridge, MA), Conseil General des Bouches du Rhone and Assistance Publique des Hopitaux de Marseille)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: Withdrawal versus RHTSH

Protocol outcome 1: Quality of life
- Actual outcome for Ablation: Physical well-being at baseline to 9 months follow up. Mean. (Mean difference (from baseline) SD: Baseline: hypothyroid: 23.8 (4.0). rhTSH: 24.0 (4.4)) 0-28 Top=High is good outcome, Comments:
(ablation period) hypothyroid:-5.78(5.68) rhTSH:-0.62(2.71)
(3 months post ablation) hypothyroid:-1.58(6.13) rhTSH:0.37(4.40)
(6 months post ablation) hypothyroid:-0.09(5.05) rhTSH:0.14(3.94)
(9 months post ablation) hypothyroid:-0.69(5.83) rhTSH:-1.11(4.86).
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: 2, Reason: loss to follow up and persistent disease. Group 2 Number missing: 1, Reason: persistent disease
- Actual outcome for Ablation: Social / familial well-being at baseline to 9 months follow up. Mean. (Mean difference (from baseline) SD: Baseline: hypothyroid: 21(4.8). rhTSH: 21.5(5.9)) 0-28 Top=High is good outcome, Comments:
(ablation period): hypothyroid:-5.0(4.18) rhTSH:-0.11(1.70)
(3 months post ablation) hypothyroid:-0.26(3.66) rhTSH:-0.32(3.18)
(6 months post ablation) hypothyroid:-0.74(6.23) rhTSH:-0.15(3.20)
(9 months post ablation) hypothyroid:0.16(4.13) rhTSH:-0.45(3.24).
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: 2, Reason: loss to follow up and persistent disease. Group 2 Number missing: 1, Reason: persistent disease
- Actual outcome for Ablation: Emotional well-being at baseline to 9 months follow up. Mean. (Mean difference (from baseline) SD: Baseline: hypothyroid: 17.6(4.2). rhTSH: 19.1(3.2)) 0-24 Top=High is good outcome, Comments:
(ablation period) hypothyroid:-0.35 (4). rhTSH: 0.86(2.39)
(3 months post ablation) hypothyroid:1.64(4.43). rhTSH:1(2.94)
(6 months post ablation) hypothyroid:0.94(5.39). rhTSH: 0.47(2.14)
(9 months post ablation) hypothyroid:1.22(4.70). rhTSH: 0.28(3.10).
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: 2, Reason: loss to follow up and persistent disease. Group 2 Number missing: 1, Reason: persistent disease

- Actual outcome for Ablation: Functional well-being at baseline to 9 months follow up. Mean. (Mean difference (from baseline) SD: Baseline: hypothyroid: 16.4(5.9). rhTSH: 18.3(5.5)) 0-28 Top=High is good outcome, Comments:
(ablation period) hypothyroid: -2.49(5.89). rhTSH: -1(3.66)
(3 months post ablation) hypothyroid: 1.77(6.46). rhTSH:0.89(3.76)
(6 months post ablation) hypothyroid: 2.12(7.14). rhTSH:1.53(3.45)
(9 months post ablation) hypothyroid: 2.19(5.37). rhTSH: 0.83(4.67).
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: 2, Reason: loss to follow up and persistent disease. Group 2 Number missing: 1, Reason: persistent disease
- Actual outcome for Ablation: Fatigue at baseline to 9 months follow up. Mean. (Mean difference (from baseline) SD: Baseline: hypothyroid: 36.1(11). rhTSH: 39.6(10.7)) 0-52 Top=High is good outcome, Comments:
(ablation period) hypothyroid:-7.31(10.35). rhTSH: -0.97(8.32)
(3 months post ablation) hypothyroid: 2.13(13.15). rhTSH: 1.14(10.26)
(6 months post ablation) hypothyroid: 2.76(13.18). rhTSH:0.75(10.87)
(9 months post ablation) hypothyroid: 3.57(13.30). rhTSH:-0.26(10.58).
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: 2, Reason: loss to follow up and persistent disease. Group 2 Number missing: 1, Reason: persistent disease
- Actual outcome for Ablation: Facit-F trial outcome index at baseline to 9 months follow up. Mean. (Mean difference (from baseline) SD: Baseline: hypothyroid: 76.1(16.8). rhTSH: 81.9(18.5)) sum of the physical well-being. functional well-being and fatigue scales 0-108 Top=High is good outcome, Comments: (ablation period) hypothyroid: -15.06(19.04). rhTSH: -2.59(12.89
(3 months post ablation) hypothyroid: 3.07(23.11). rhTSH:2.40(16.40)
(6 months post ablation) hypothyroid: 5.18 (22.12). rhTSH:2.42 (16.26)
(9 months post ablation) hypothyroid: 5.30(21.21). rhTSH: -0.51(18.60).
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: 2, Reason: loss to follow up and persistent disease. Group 2 Number missing: 1, Reason: persistent disease
- Actual outcome for Ablation: Facit-G total score at baseline to 9 months follow up. Mean. (Mean difference (from baseline) SD: Baseline: hypothyroid: 78.6(13.3). rhTSH: 84.4(16.6)) sum of physical, social, emotional and functional wellbeing scores 0-108 Top=High is good outcome, Comments:
(ablation period) hypothyroid:-9.82(16.87). rhTSH: 1.63(7.72)
(3 months post ablation) hypothyroid:2.83 (14.89). rhTSH: 2.37(9.83)
(6 months post ablation) hypothyroid:1.82(18.80). rhTSH: 1.85(8.22)
(9 months post ablation) hypothyroid:4.80(13.15). rhTSH: -0.10(10.82).
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: 2, Reason: loss to follow up and persistent disease. Group 2 Number missing: 1, Reason: persistent disease
- Actual outcome for Ablation: Facit-F total score at baseline to 9 months follow up. Mean. (Mean difference (from baseline) SD: Baseline: hypothyroid: 114.2(20.2). rhTSH: 125.1(24.6)) sum of FACT-G score and Fatigue subscale, Comments:

t1 (ablation period) hypothyroid:-16.26(25.96) rhTSH:-4.05(15.83) t2 (3 months post ablation) hypothyroid:7.77(23.92) rhTSH:4.26(18.91) t3 (6 months post ablation) hypothyroid:5.28(26.32) rhTSH:1.40(18.24) t4 (9 months post ablation) hypothyroid:11.13(22.77) rhTSH:0.80(20.35)). 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: 2, Reason: loss to follow up and persistent disease. Group 2 Number missing: 1, Reason: persistent disease Protocol outcome 2: Incidence of distant metastases - Actual outcome for Ablation: Metastatic lymph nodes at 9 months post radioiodine treatment. Group 1: 1/35, Group 2: 0/36. Comments: p value 0.49 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: 2, Reason: loss to follow up and persistent disease. Group 2 Number missing: 1, Reason: persistent disease Protocol outcome 3: Successful ablation - Actual outcome for Ablation: successful ablation at 9 months post radioiodine treatment. Group 1: 34/35, Group 2: 32/36. Comments: successful ablation considered as Thyroglobulin < 0.8µg/l and an uptake of <0.1% on diagnostic whole body scan 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: loss to follow up and persistent disease. Group 2 Number missing: 1, Reason: persistent disease
Protocol outcomes not reported by the study	Mortality . Local cancer progression . Cancer recurrence

Appendix E – Forest plots

E.1 Radioiodine ablation with withdrawal of levothyroxine to radioiodine ablation with thyrotropin alfa

Figure 5: Successful ablation (Tg <0.2ng/ml)

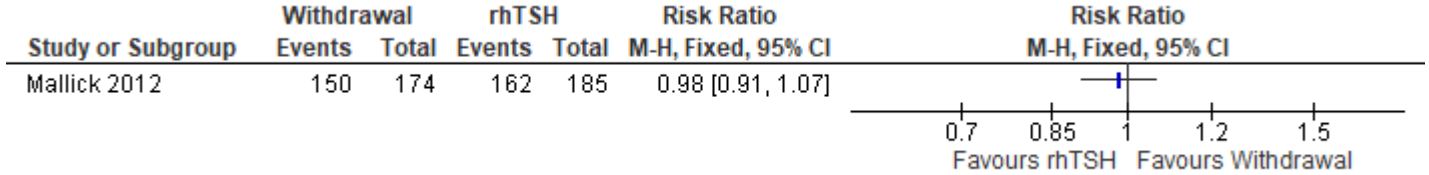


Figure 6: Successful ablation (Tg <0.2ng/ml) and <0.1% WBS

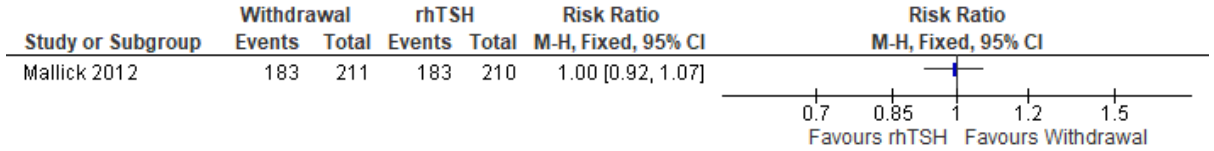


Figure 7: Successful ablation (Tg <1ng/ml)

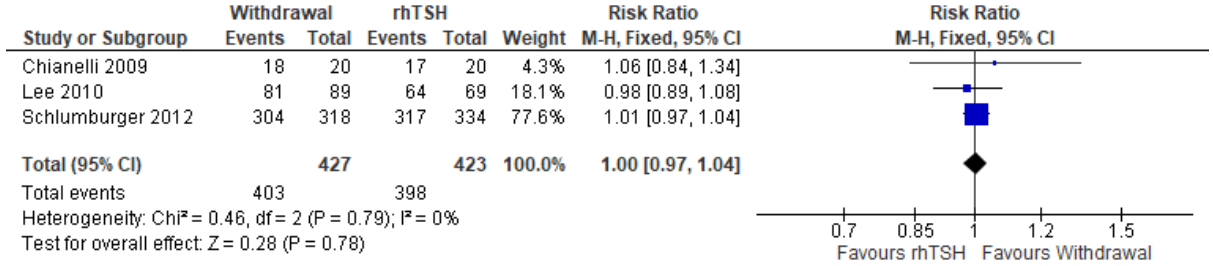


Figure 8: Successful uptake (no visible uptake)

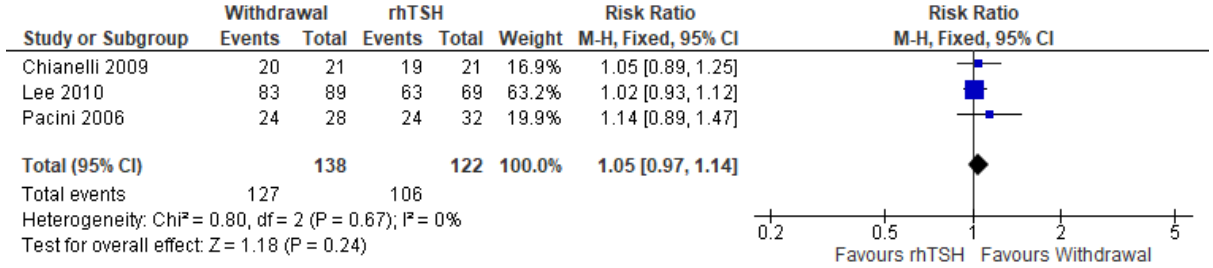


Figure 9: Successful ablation (Tg<0.8µg/l + <0.1% WBS)

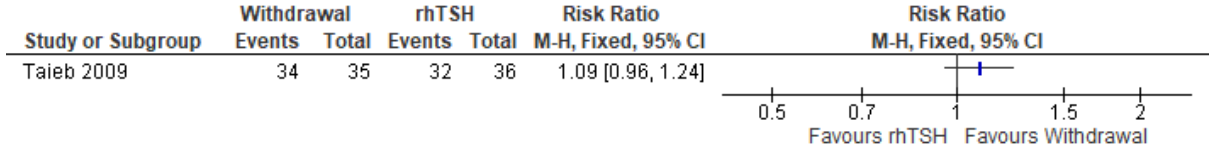


Figure 10: Complete ablation

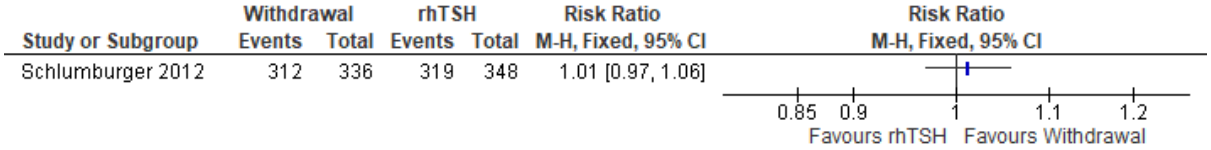


Figure 11: Visible uptake <0.1%

Figure 12: Lymph node metastases

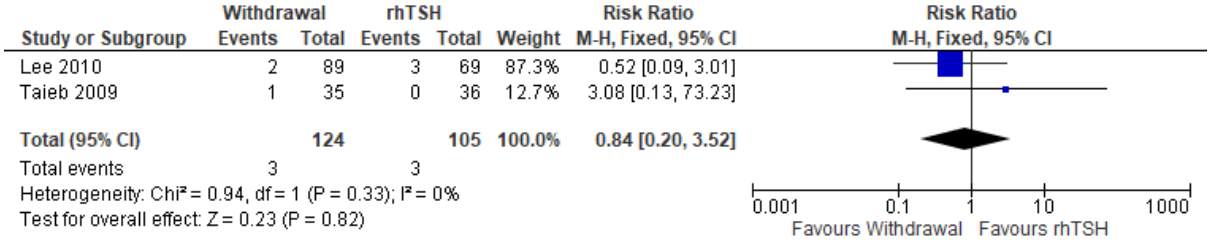


Figure 13: Cancer recurrence

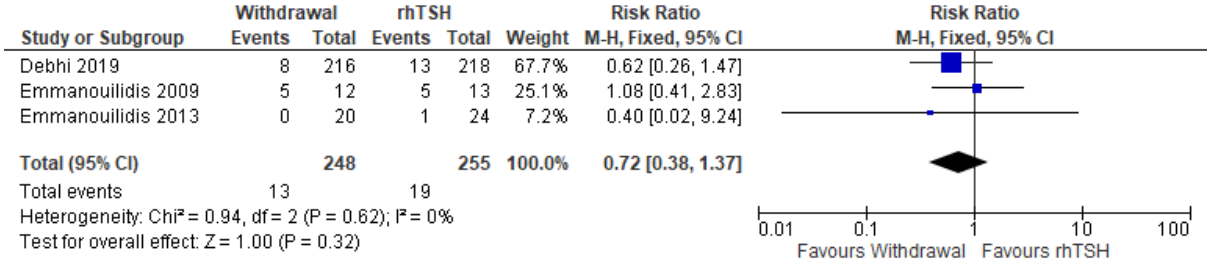


Figure 14: Thyroglobulin levels ng/ml (12 months - 2.5 years)

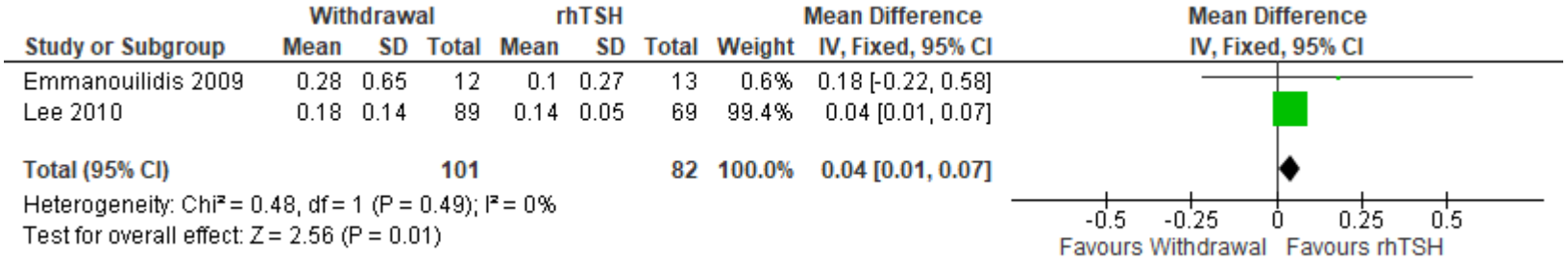


Figure 15: SF-36 score (mental component)

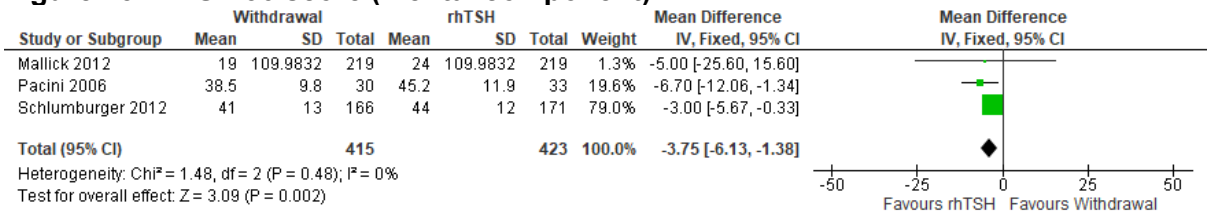
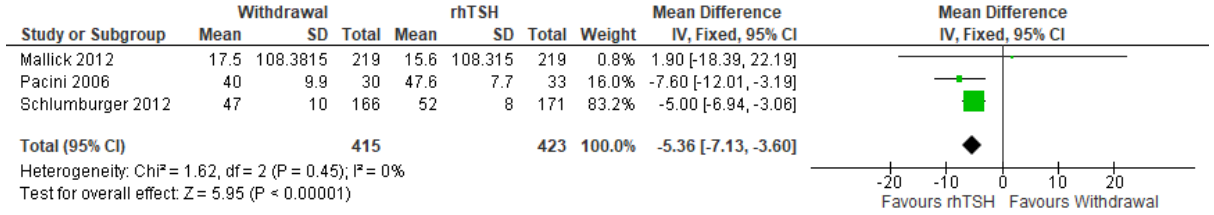
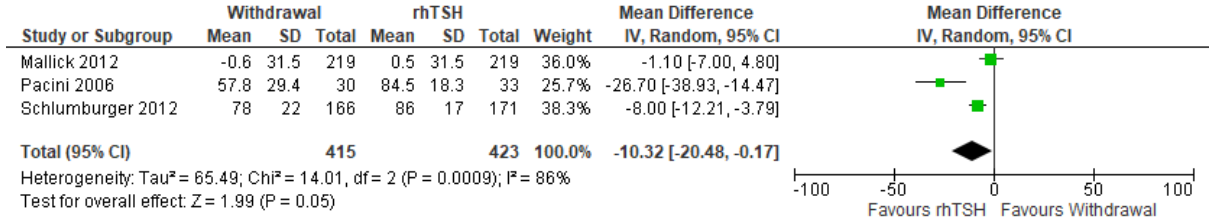


Figure 16: SF-36 score (physical component)



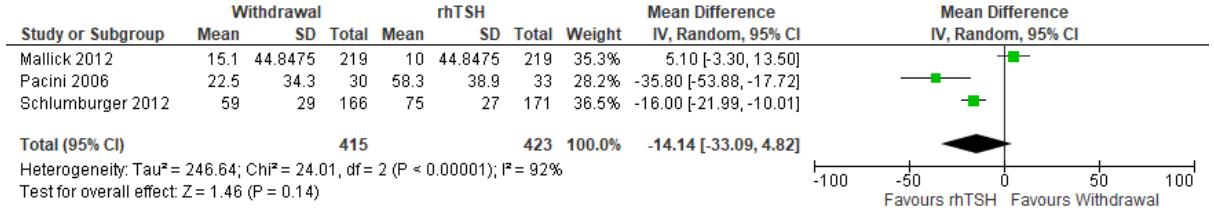
1

Figure 17: SF-36 physical functioning score



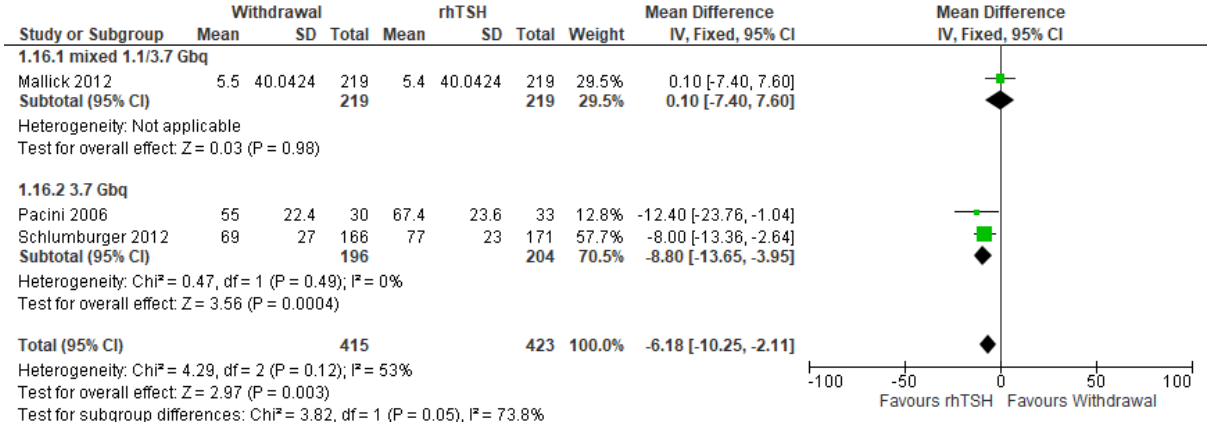
2

Figure 18: SF-36 role – physical score



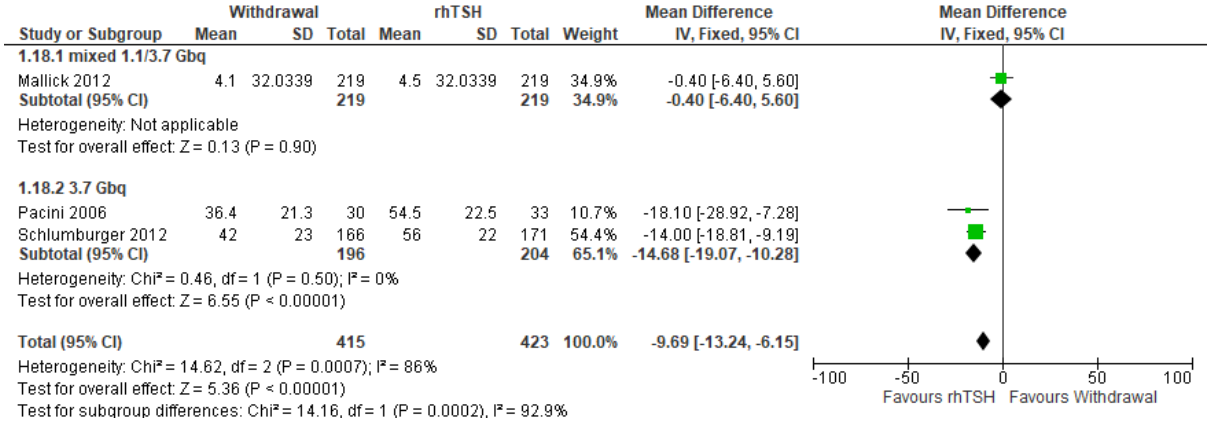
3

Figure 19: SF-35 bodily pain score



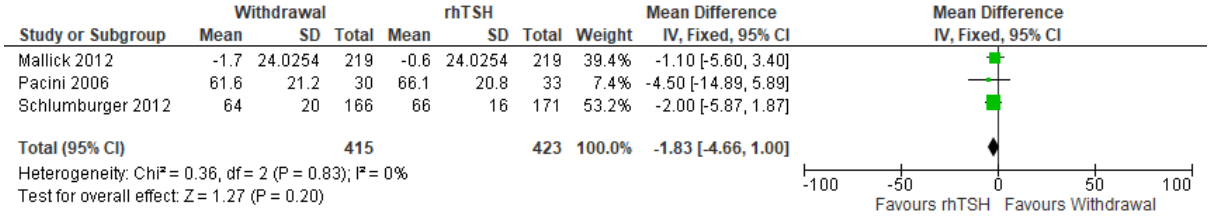
1

Figure 20: SF-36 vitality score



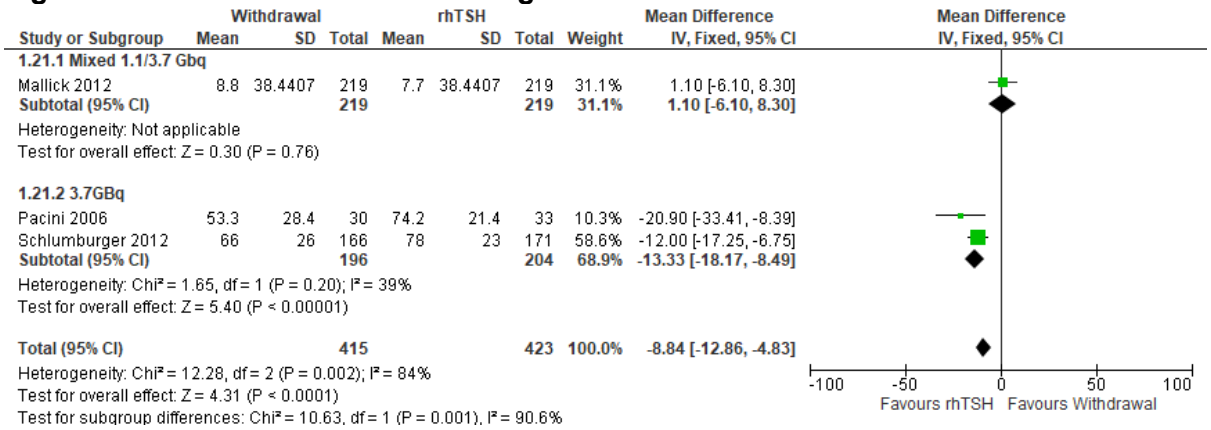
2

Figure 21: SF-36 general health score



1

Figure 22: SF-36 social functioning score



2

Figure 23: SF-36 role – emotional score

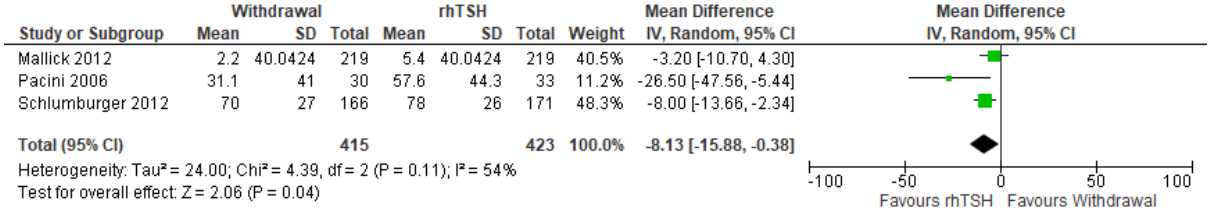


Figure 24: SF-36 mental health score

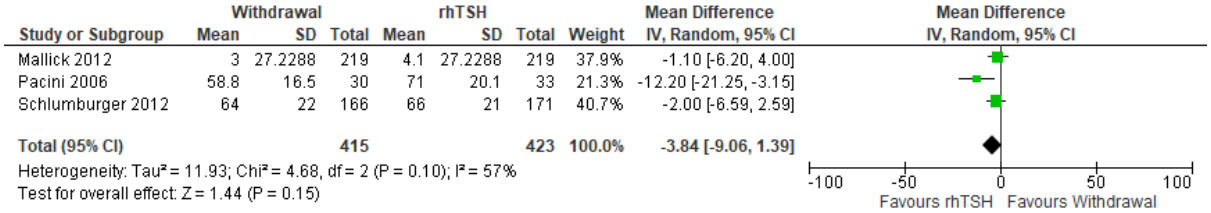


Figure 25: EQ5D Utility score

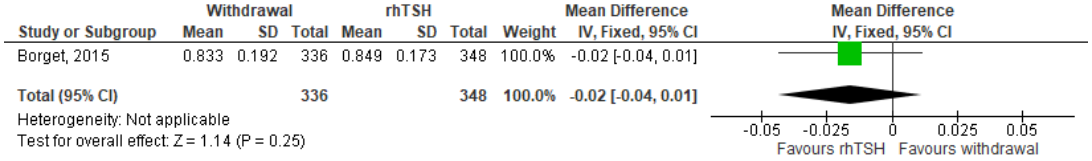


Figure 26: Physical well-being (ablation period)

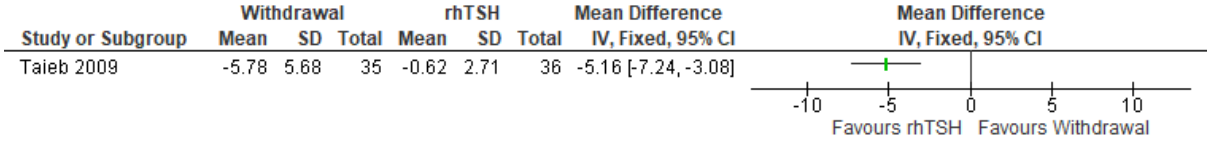


Figure 27: Physical well-being (3 months post ablation)

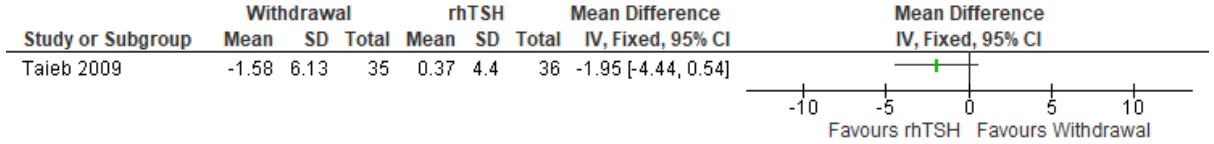


Figure 28: Physical well-being (6 months post ablation)

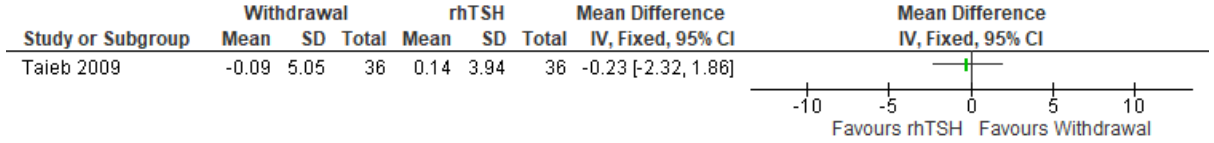


Figure 29: Physical well-being (9 months post ablation)

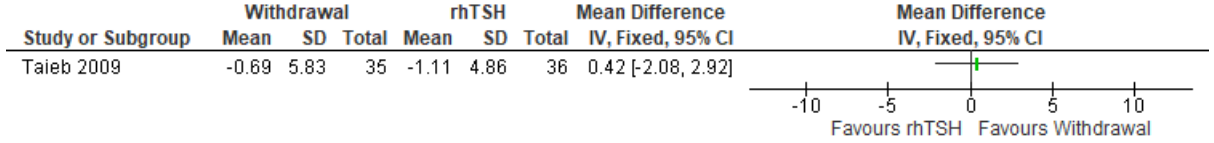


Figure 30: Social / familial well-being (ablation period)

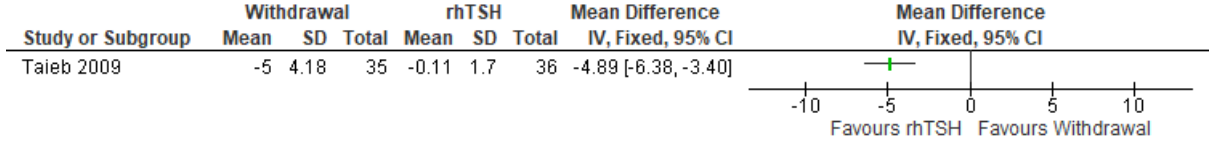


Figure 31: Social / familial well-being (3 months post ablation)

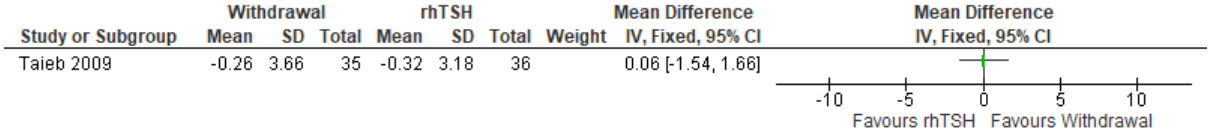


Figure 32: Social / familial well-being (6 months post ablation)

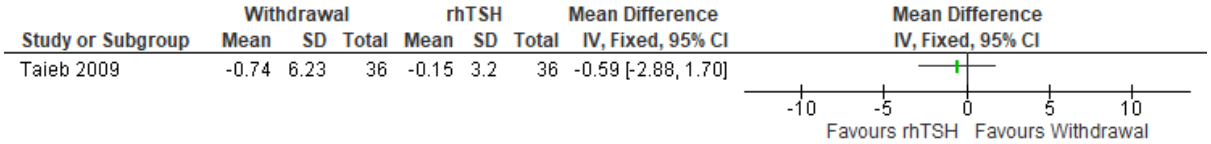


Figure 33: Social / familial well-being (9 months post ablation)

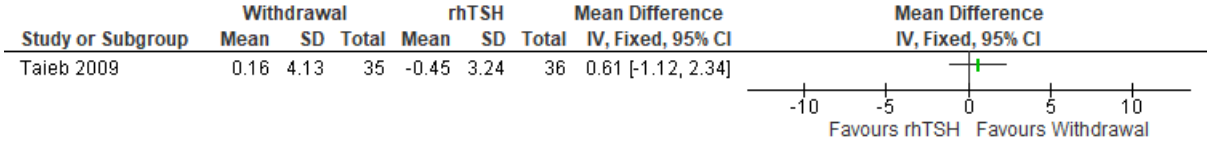


Figure 34: Emotional well-being (ablation period)

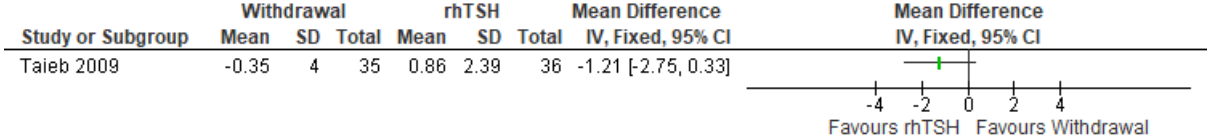


Figure 35: Emotional well-being (3 months post ablation)

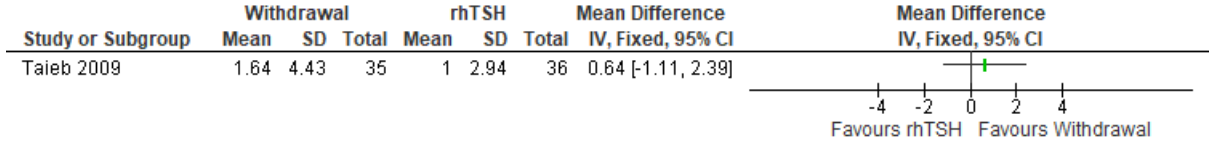


Figure 36: Emotional well-being (6 months post ablation)

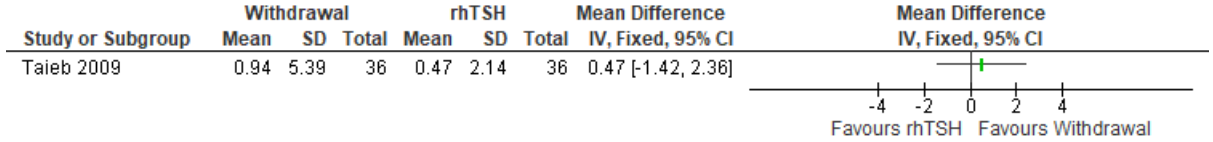


Figure 37: Emotional well-being (9 months post ablation)

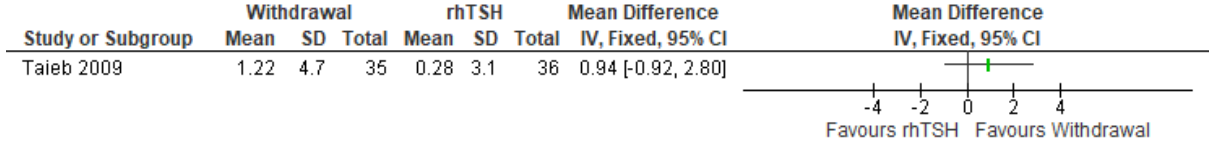


Figure 38: Functional well-being (ablation period)

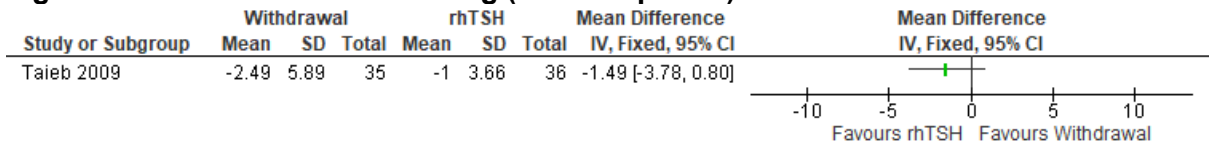


Figure 39: Functional well-being (3 months post ablation)

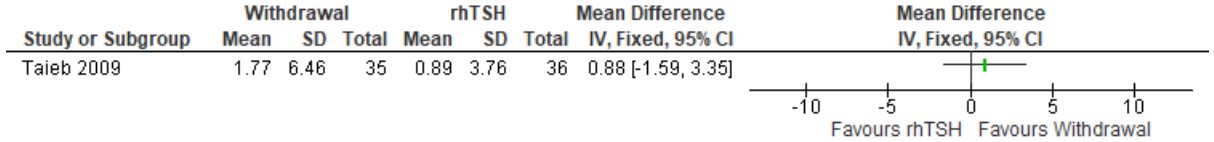


Figure 40: Functional well-being (6 months post ablation)

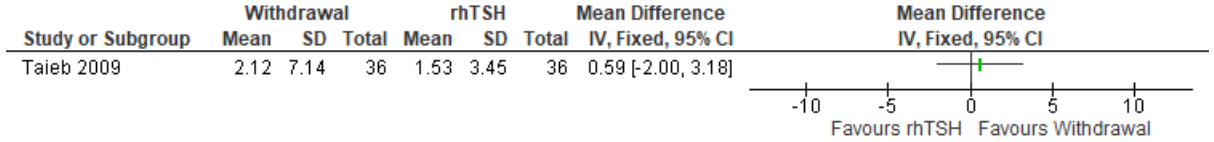


Figure 41: Functional well-being (9 months post ablation)

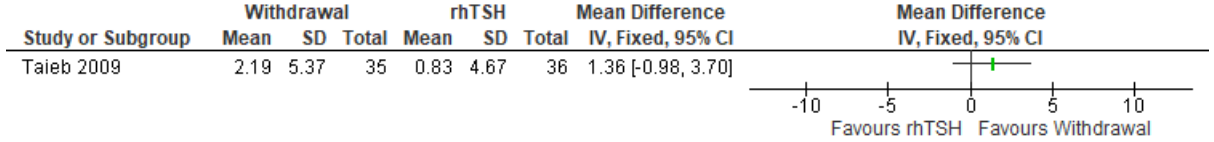


Figure 42: Fatigue (ablation period)

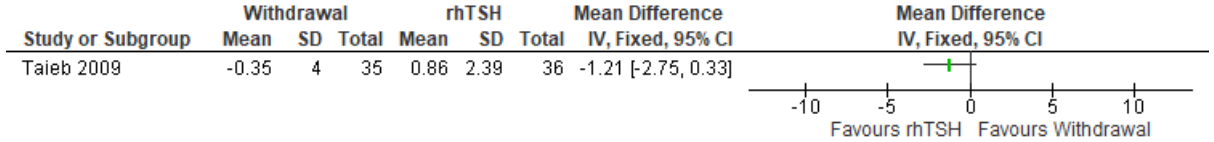


Figure 43: Fatigue (3 months post ablation)

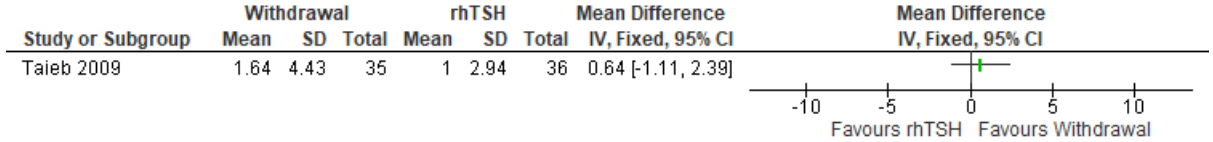


Figure 44: Fatigue (6 months post ablation)

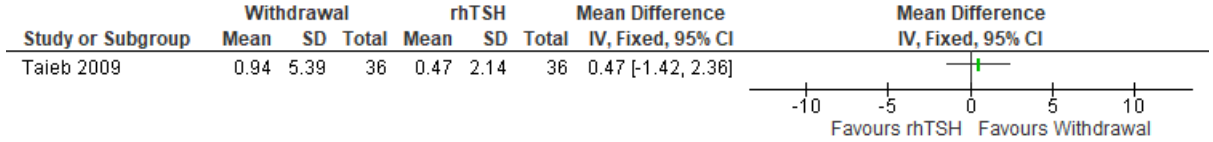


Figure 45: Fatigue (9 months post ablation)

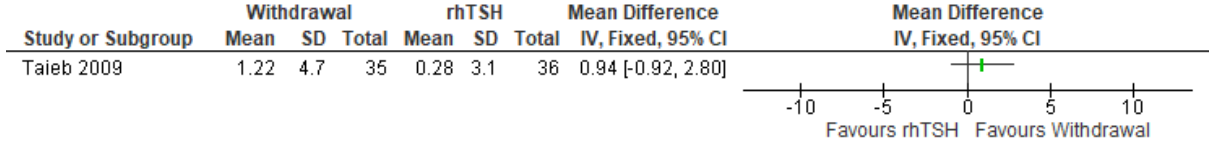


Figure 46: FACIT-F (ablation period)

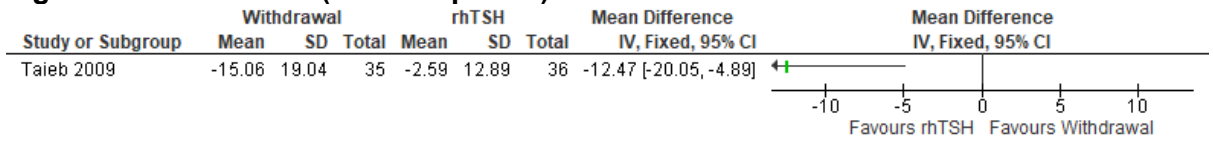


Figure 47: FACIT-F (3 months post ablation)

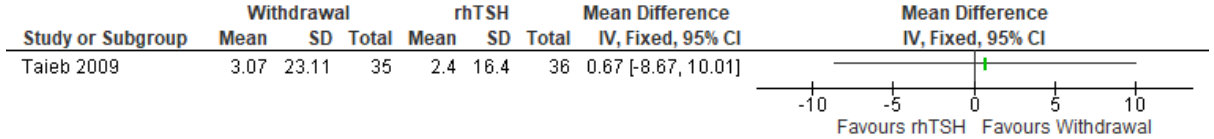


Figure 48: FACIT-F (6 months post ablation)

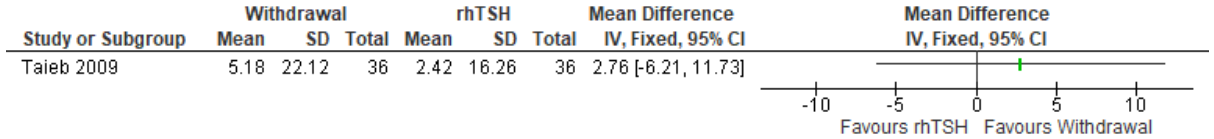


Figure 49: FACIT-F (9 months post ablation)

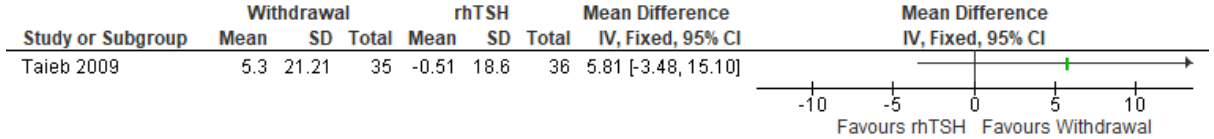


Figure 50: FACT-G (total score) ablation period

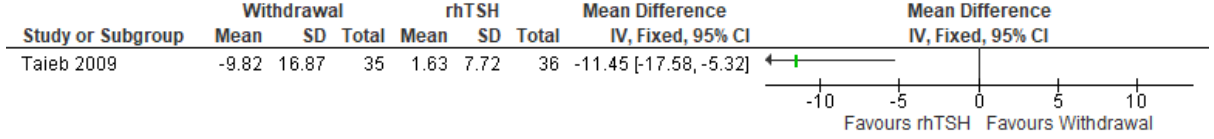


Figure 51: FACT-G (total score) 3 months post ablation

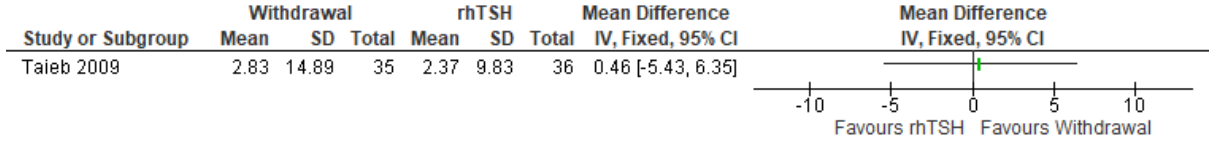


Figure 52: FACT-G (total score) 6 months post ablation

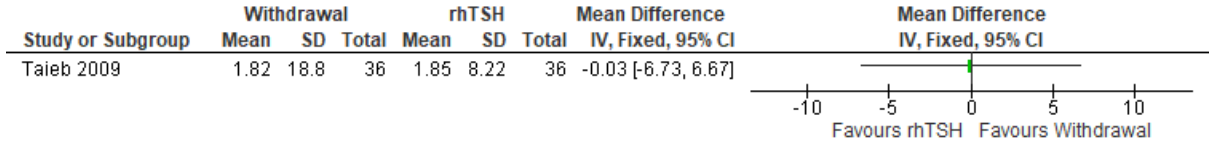


Figure 53: FACT-G (total score) 9 months post ablation

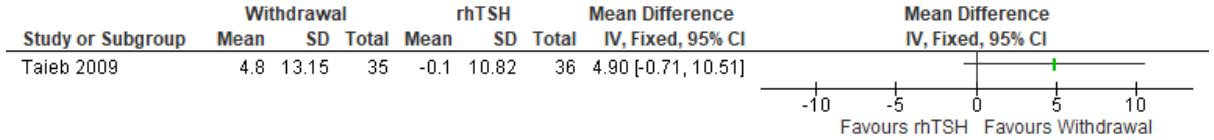


Figure 54: FACIT-F (total score) ablation period

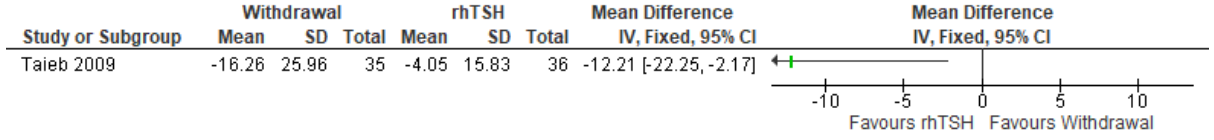


Figure 55: FACIT-F (total score) 3 months post ablation period

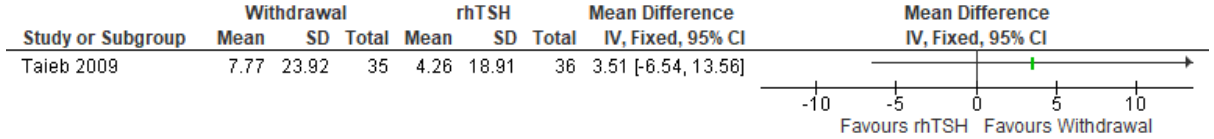


Figure 56: FACIT-F (total score) 6 months post ablation period

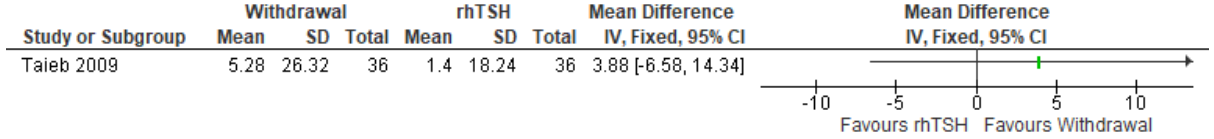
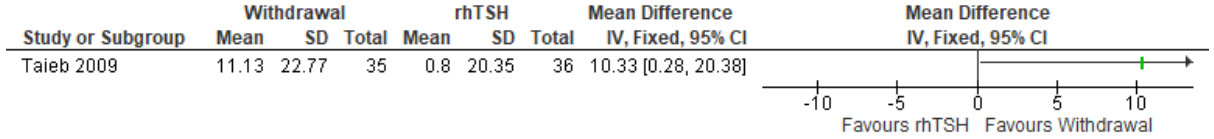


Figure 57: FACIT-F (total score) 9 months post ablation period



Appendix F – GRADE tables

Table 10: Clinical evidence profile: Radioiodine ablation with prior withdrawal of thyrotropin alfa or with thyrotropin alfa

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Withdrawal	RhTSH	Relative (95% CI)	Absolute		
Successful ablation (Tg<0.2ng/ml) (follow-up 3 months)												
1	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	150/174 (86.2%)	87.6%	RR 0.98 (0.91 to 1.07)	18 fewer per 1000 (from 79 fewer to 61 more)	⊕⊕OO LOW	CRITICAL
Successful ablation (Tg<0.2ng/ml) and <0.1 WBS% (follow-up 3 months)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	183/211 (86.7%)	87.1%	RR 1 (0.92 to 1.07)	0 fewer per 1000 (from 70 fewer to 61 more)	⊕⊕⊕O MODERATE	CRITICAL
Successful ablation (Tg<1ng/ml) (follow-up 6-9 months)												
3	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	403/427 (94.4%)	94.1%	RR 1 (0.97 to 1.04)	0 fewer per 1000 (from 28 fewer to 37 more)	⊕⊕⊕O MODERATE	CRITICAL
Successful ablation (no visible uptake) (follow-up 6-12 months)												
3	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	127/138 (92%)	90.5%	RR 1.05 (0.97 to 1.14)	45 more per 1000 (from 27 fewer to 127 more)	⊕⊕⊕⊕ HIGH	CRITICAL
Successful ablation (Tg<0.8Âµg/l + <0.1% WBS uptake) (follow-up 9 months)												
1	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	34/35 (97.1%)	88.9%	RR 1.09 (0.96 to 1.24)	80 more per 1000 (from 36 fewer to 213 more)	⊕⊕⊕⊕ HIGH	CRITICAL
Complete Ablation (follow-up 6-10 months)												
1	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	312/336 (92.9%)	91.7%	RR 1.01 (0.97 to 1.06)	9 more per 1000 (from 28 fewer to 55 more)	⊕⊕OO LOW	CRITICAL

Visible uptake <0.1% (follow-up 6-9 months)												
2	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	202/239 (84.5%)	59.4%	RR 0.98 (0.93 to 1.04)	12 fewer per 1000 (from 42 fewer to 24 more)	⊕⊕⊕⊕ HIGH	CRITICAL
Lymph node metastases (follow-up 9-12 months)												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	3/124 (2.4%)	2.2%	RR 0.84 (0.2 to 3.52)	4 fewer per 1000 (from 18 fewer to 55 more)	⊕OOO VERY LOW	CRITICAL
Cancer recurrence (follow-up up to 4.5 years)												
3	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[2]	none	13/248 (5.2%)	6%	RR 0.72 (0.38 to 1.37)	17 fewer per 1000 (from 37 fewer to 22 more)	⊕⊕OO LOW	CRITICAL
Thyroglobulin levels (ng/ml) (follow-up 12 months - 2.5 years. Better indicated by lower values)												
2	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision[2]	none	101	82	-	MD 0.04 higher (0.01 to 0.07 higher)	⊕⊕⊕⊕ HIGH	CRITICAL
SF-36 score (mental component) (follow-up 1-4 months. range of scores: 0-100. Better indicated by higher values)												
3	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious imprecision[2]	none	415	423	-	MD 3.75 lower (6.13 to 1.38 lower)	⊕⊕OO LOW	CRITICAL
SF-36 score (physical component) (follow-up 1-4 months. range of scores: 0-100. Better indicated by higher values)												
3	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious imprecision[2]	none	415	423	-	MD 5.36 lower (7.13 to 3.60 lower)	⊕⊕OO LOW	CRITICAL
SF-36 (physical functioning score) (follow-up 1-4 months. range of scores: 0-100. Better indicated by higher values)												
3	randomised trials	serious[1]	Very serious inconsistency[3]	no serious indirectness	serious imprecision[2]	none	415	423	-	MD 10.32 lower (20.48 to 0.17 lower)	⊕⊖⊖⊖ VERY LOW	CRITICAL
SF-36 (role physical) (follow-up 1-4 months. range of scores: 0-100. Better indicated by higher values)												
3	randomised trials	serious[1]	Very serious inconsistency[3]	no serious indirectness	serious imprecision[2]	none	415	423	-	MD 14.14 lower (33.09 lower to 4.82 higher)	⊕⊖⊖⊖ VERY LOW	CRITICAL
SF-36 (bodily pain) (follow-up 4 months. range of scores: 0-100. Better indicated by higher values) SUBGROUPED TO MIXED 1.1/3.7 Gbq												

1	randomised trials	serious[1]	NA	no serious indirectness	no serious imprecision	none	219	219	-	MD 0.10 higher (7.40 lower to 7.60 higher)	⊕⊕⊕O MODERATE	CRITICAL
SF-36 (bodily pain) (follow-up 1 month. range of scores: 0-100. Better indicated by higher values) SUBGROUPED TO 3.7 Gbq												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious imprecision[2]	none	196	204	-	MD 8.80 lower (13.65 lower to 3.95 lower)	⊕⊕OO LOW	CRITICAL
SF-36 (vitality) (follow-up 1-4 months. range of scores: 0-100. Better indicated by higher values) SUBGROUPED TO MIXED 1.1/3.7 Gbq												
1	randomised trials	serious[1]	NA	no serious indirectness	no serious imprecision	none	219	219	-	MD 0.40 lower (6.40 lower to 5.60 higher)	⊕⊕⊕O MODERATE	CRITICAL
SF-36 (vitality) (follow-up 1 month. range of scores: 0-100. Better indicated by higher values) SUBGROUPED TO 3.7 Gbq												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious imprecision[2]	none	196	204	-	MD 14.68 lower (19.07 lower to 10.28 lower)	⊕⊕OO LOW	CRITICAL
SF-36 (general health) (follow-up 1-3 months. range of scores: 0-100. Better indicated by higher values)												
3	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	415	423	-	MD 1.83 lower (4.66 lower to 1.00 higher)	⊕⊕⊕O MODERATE	CRITICAL
SF-36 (social functioning score) (follow-up 1-4 months. range of scores: 0-100. Better indicated by higher values) SUBGROUPED TO MIXED 1.1/3.7 Gbq												
1	randomised trials	serious[1]	NA	no serious indirectness	no serious imprecision	none	219	219	-	MD 1.10 higher (6.10 lower to 8.30 higher)	⊕⊕⊕O MODERATE	CRITICAL
SF-36 (social functioning score) (follow-up 1 month. range of scores: 0-100. Better indicated by higher values) SUBGROUPED TO 3.7 Gbq												
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious imprecision[2]	none	196	204	-	MD 13.33 lower (18.17 lower to 8.49 lower)	⊕⊕OO LOW	CRITICAL
SF-36 (role - emotional score) (follow-up 1-3 months. range of scores: 0-100. Better indicated by higher values)												
3	randomised trials	serious[1]	Serious inconsistency[3]	no serious indirectness	no serious imprecision	none	415	423	-	MD 8.13 lower (15.88 lower to 0.38 lower)	⊕⊕⊖⊖ LOW	CRITICAL
SF-36 (mental health score) (follow-up 1-3 months. range of scores: 0-100. Better indicated by higher values)												
3	randomised trials	serious[1]	Serious inconsistency[3]	no serious indirectness	no serious imprecision	none	415	423	-	MD 3.84 lower (9.06 lower to 1.39 higher)	⊕⊕⊖⊖ LOW	CRITICAL

EQ5D Utility score (follow up 8 months. range of scores 0-1. Better indicated by higher values)												
1	randomised trials	serious[1]	NA	no serious indirectness	no serious imprecision	none	336	348	-	MD 0.02 lower (0.04 lower to 0.01 higher)	⊕⊕⊕⊖ MODERATE	CRITICAL
Physical Well-being (follow-up ablation period. range of scores: 0-28. Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	35	36	-	MD 5.16 lower (7.24 to 3.08 lower)	⊕⊕⊕O MODERATE	CRITICAL
Physical Well-being (follow-up 3 months post ablation. range of scores: 0-28. Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	35	36	-	MD 1.95 lower (4.44 lower to 0.54 higher)	⊕⊕OO LOW	CRITICAL
Physical Well-being (follow-up 6 months post ablation. range of scores: 0-28. Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	36	36	-	MD 0.23 lower (2.32 lower to 1.86 higher)	⊕⊕OO LOW	CRITICAL
Physical Well-being (follow-up 9 months post ablation. range of scores: 0-28. Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	35	36	-	MD 0.42 higher (2.08 lower to 2.92 higher)	⊕⊕OO LOW	CRITICAL
Social / Familial Well-being (follow-up ablation period. range of scores: 0-28. Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	35	36	-	MD 4.89 lower (6.38 to 3.4 lower)	⊕⊕⊕O MODERATE	CRITICAL
Social / Familial Well-being (follow-up 3 months post ablation period. range of scores: 0-28. Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	35	36	-	MD 0.06 higher (1.54 lower to 1.66 higher)	⊕⊕OO LOW	CRITICAL
Social / Familial Well-being (follow-up 6 months post ablation. range of scores: 0-28. Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	36	36	-	MD 0.59 lower (2.88 lower to 1.7 higher)	⊕OOO VERY LOW	CRITICAL
Social / Familial Well-being (follow-up 9 months post ablation. range of scores: 0-28. Better indicated by higher values)												

1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	35	36	-	MD 0.61 higher (1.12 lower to 2.34 higher)	⊕⊕OO LOW	CRITICAL
Emotional Well-being (follow-up ablation period. range of scores: 0-24. Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	35	36	-	MD 1.21 lower (2.75 lower to 0.33 higher)	⊕⊕OO LOW	CRITICAL
Emotional Well-being (follow-up 3 months post ablation. range of scores: 0-24. Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	35	36	-	MD 0.64 higher (1.11 lower to 2.39 higher)	⊕⊕OO LOW	CRITICAL
Emotional Well-being (follow-up 6 months post ablation. range of scores: 0-24. Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	36	36	-	MD 0.47 higher (1.42 lower to 2.36 higher)	⊕OOO VERY LOW	CRITICAL
Emotional Well-being (follow-up 9 months post ablation. range of scores: 0-24. Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	35	36	-	MD 0.94 higher (0.92 lower to 2.8 higher)	⊕⊕OO LOW	CRITICAL
Functional Well-being (follow-up ablation period. range of scores: 0-28. Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	35	36	-	MD 1.49 lower (3.78 lower to 0.8 higher)	⊕⊕OO LOW	CRITICAL
Functional Well-being (follow-up 3 months post ablation. range of scores: 0-28. Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	35	36	-	MD 0.88 higher (1.59 lower to 3.35 higher)	⊕⊕OO LOW	CRITICAL
Functional Well-being (follow-up 6 months post ablation. range of scores: 0-28. Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	36	36	-	MD 0.59 higher (2 lower to 3.18 higher)	⊕OOO VERY LOW	CRITICAL
Functional Well-being (follow-up 9 months post ablation. range of scores: 0-28. Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	35	36	-	MD 1.36 higher (0.98 lower to 3.7 higher)	⊕⊕OO LOW	CRITICAL

Fatigue (follow-up ablation period. range of scores: 0-52. Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	35	36	-	MD 1.21 lower (2.75 lower to 0.33 higher)	⊕⊕OO LOW	CRITICAL
Fatigue (follow-up 3 months post ablation. range of scores: 0-52. Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	35	36	-	MD 0.64 higher (1.11 lower to 2.39 higher)	⊕⊕OO LOW	CRITICAL
Fatigue (follow-up 6 months post ablation. range of scores: 0-52. Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	36	36	-	MD 0.47 higher (1.42 lower to 2.36 higher)	⊕OOO VERY LOW	CRITICAL
Fatigue (follow-up 9 months post ablation. range of scores: 0-52. Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	35	36	-	MD 0.94 higher (0.92 lower to 2.8 higher)	⊕⊕OO LOW	CRITICAL
Facit-F (TOI) (follow-up ablation period. range of scores: 0-52. Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	35	36	-	MD 12.47 lower (20.05 to 4.89 lower)	⊕⊕OO LOW	CRITICAL
Facit-F (TOI) (follow-up 3 months post ablation. range of scores: 0-108. Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	35	36	-	MD 0.67 higher (8.67 lower to 10.01 higher)	⊕OOO VERY LOW	CRITICAL
Facit-F (TOI) (follow-up 6 months post ablation. range of scores: 0-108. Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	36	36	-	MD 2.76 higher (6.21 lower to 11.73 higher)	⊕⊕OO LOW	CRITICAL
Facit-F (TOI) (follow-up 9 months post ablation. range of scores: 0-108. Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	35	36	-	MD 5.81 higher (3.48 lower to 15.1 higher)	⊕⊕OO LOW	CRITICAL
FACT-G (total score) (follow-up ablation period. range of scores: 0-108. Better indicated by higher values)												

1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	35	36	-	MD 11.45 lower (17.58 to 5.32 lower)	⊕⊕⊕O MODERATE	CRITICAL
FACT-G (total score) (follow-up 3 months post ablation. range of scores: 0-108. Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	35	36	-	MD 0.46 higher (5.43 lower to 6.35 higher)	⊕⊕OO LOW	CRITICAL
FACT-G (total score) (follow-up 6 months post ablation. range of scores: 0-108. Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	none	36	36	-	MD 0.03 lower (6.73 lower to 6.67 higher)	⊕OOO VERY LOW	CRITICAL
FACT-G (total score) (follow-up 9 months post ablation. range of scores: 0-108. Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	35	36	-	MD 4.9 higher (0.71 lower to 10.51 higher)	⊕⊕OO LOW	CRITICAL
Facit-F (total score) (follow-up ablation period. range of scores: 0-160. Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	35	36	-	MD 12.21 lower (22.25 to 2.17 lower)	⊕⊕OO LOW	CRITICAL
Facit-F (total score) (follow-up 3 months post ablation period. range of scores: 0-160. Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	35	36	-	MD 3.51 higher (6.54 lower to 13.56 higher)	⊕⊕OO LOW	CRITICAL
Facit-F (total score) (follow-up 6 months post ablation. range of scores: 0-160. Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	36	36	-	MD 3.88 higher (6.58 lower to 14.34 higher)	⊕⊕OO LOW	CRITICAL
Facit-F (total score) (follow-up 9 months post ablation. range of scores: 0-160. Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	35	36	-	MD 10.33 higher (0.28 to 20.38 higher)	⊕⊕OO LOW	CRITICAL

[1] Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
[2] Downgraded by 1 increment if the confidence interval crossed one MID* or by 2 increments if the confidence interval crossed both MIDs
[3] Downgraded by 1 increment if the I^2 was between 50% and 75% by 2 increments if the I^2 was over 75%

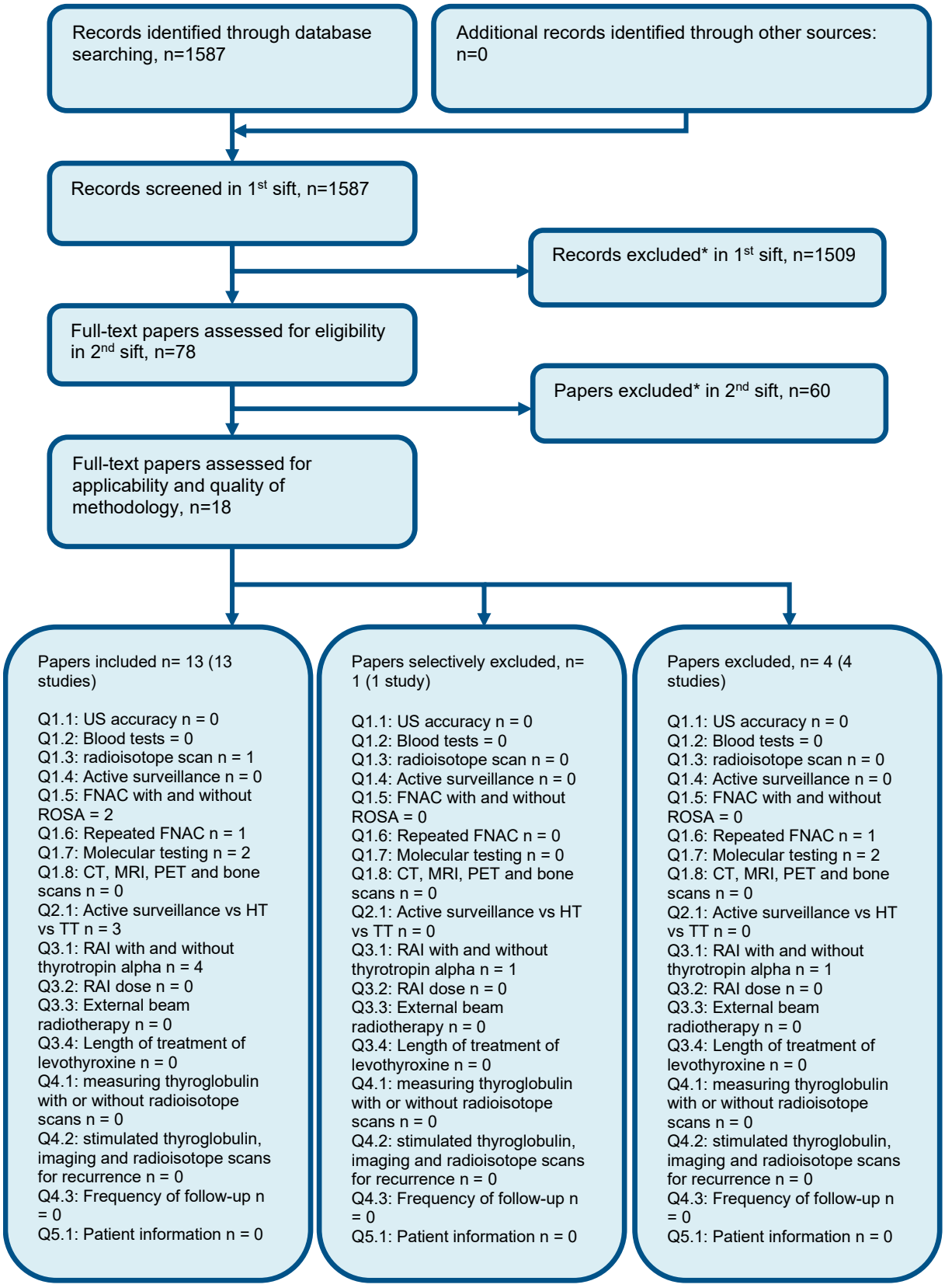
*The MIDs for binary outcomes were based on default OR, RR or HR values of 0.8 or 1.25. For continuous variables, the MIDs were based on the default value of ±0.5 x the median standard deviation (sd) in the control group. The median control group sd, together with the MID for all continuous variables, have been tabulated below:

Outcome	Control group median sd	MID
Thyroglobulin levels (ng/ml)	0.16	0.08
SF-36 score (mental component)	12	6
SF-36 score (physical component)	8	4
SF-36 (physical functioning score)	18.3	9.15
SF-36 (role physical)	38.9	19.45
SF-36 (bodily pain) SUBGROUPED TO MIXED 1.1/3.7 Gbq	40.04	20.02
SF-36 (bodily pain) SUBGROUPED TO 3.7 Gbq	23.3	11.65
SF-36 (vitality) SUBGROUPED TO MIXED 1.1/3.7 Gbq	32.02	16.01
SF-36 (vitality) SUBGROUPED TO 3.7 Gbq	22.25	11.12
SF-36 (general health)	20.8	10.4
SF-36 (social functioning score) SUBGROUPED TO MIXED 1.1/3.7 Gbq	38.44	19.22
SF-36 (social functioning score) SUBGROUPED TO 3.7 Gbq	22.2	11.1
SF-36 (role - emotional score)	40.04	20.02
SF-36 (mental health score)	21	10.5
EQ5D Utility score:	0.173	0.0865
Physical Well-being 0	2.71	1.35
Physical Well-being 3m	4.4	2.2
Physical Well-being 6m	3.94	1.97
Physical Well-being 9m	4.86	2.43
Social / Familial Well-being 0	1.7	0.85
Social / Familial Well-being 3m	3.18	1.59
Social / Familial Well-being 6m	3.2	1.6
Social / Familial Well-being 9m	3.24	1.62
Emotional Well-being 0	2.39	1.2
Emotional Well-being 3m	2.94	1.47
Emotional Well-being 6m	2.14	1.07
Emotional Well-being 9m	3.1	1.55
Functional Well-being 0	3.66	1.83
Functional Well-being 3m	3.76	1.88
Functional Well-being 6m	3.45	1.73
Functional Well-being 9m	4.67	2.3
Fatigue 0	2.39	1.2
Fatigue 3m	2.94	1.5
Fatigue 6m	2.14	1.07
Fatigue 9m	3.1	1.55
Facit-F (TOI) 0	12.89	6.45
Facit-F (TOI) 3m	16.4	8.2
Facit-F (TOI) 6m	16.26	8.13
Facit-F (TOI) 9m	18.6	9.3
FACT-G (total score) 0	7.72	3.86
FACT-G (total score) 3m	9.83	4.93
FACT-G (total score) 6m	8.22	4.11
FACT-G (total score) 9m	10.82	5.41

Facit-F (total score) 0	15.83	7.92
Facit-F (total score) 3m	18.91	9.45
Facit-F (total score) 6m	18.24	9.12
Facit-F (total score) 9m	20.35	10.17

1

Appendix G – Economic evidence study selection



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

Appendix H – Economic evidence tables

Study	Borget 2015 [7]			
Study details	**Population & interventions**	**Costs**	**Health outcomes**	**Cost effectiveness**
Economic analysis: CUA (health outcome: QALYs) **Study design:** Randomized trial with 2-by-2 design **Approach to analysis:** Within-trial CEA. mean costs were compared using non-parametric Wilcoxon tests **Perspective:** French societal perspective **Follow-up:** 8 months **Treatment effect duration:**[(a)] 8 months **Discounting:** Costs: NA Outcomes: NA	**Population:** Adults (≥18 years) who underwent total thyroidectomy for low-risk differentiated thyroid cancer and were receiving TSH stimulation in preparation for post-thyroidectomy radioiodine ablation **Cohort settings:** Start age: NR Male: NR N: 684 evaluable patients **Intervention 1:** Endogenous stimulation of TSH with THW prior to radioiodine ablation **Intervention 2:** Exogenous stimulation of TSH with rhTSH prior to radioiodine ablation	**Total costs (mean per patient):** Intervention 1: £2,342 Intervention 2: £ £2,924 Incremental (2−1): £582 (95% CI: £523 to £641. p=NR) **Currency & cost year:** 2013 French euros (presented here as 2013 UK pounds[(b)]) **Cost components incorporated:** Intervention cost, fixed hospital costs (staff, equipment, overhead), variable hospital costs (resources required for radioiodine administration, rhTSH, radioiodine activity).	**QALYs (mean per patient):** Intervention 1: 0.675 Intervention 2: 0.687 Incremental (2−1): 0.012 (95% CI: -0.002 to 0.028. p=NR)	**ICER (Intervention 2 versus Intervention 1):** £48,500 per QALY gained (da) 95% CI: NR Probability that Intervention 2 was cost effective (£20K/30K threshold): 1.5%/22% **Analysis of uncertainty:** When the cost of rhTSH was reduced by 30%, the probability that rhTSH was cost effective at a threshold of £42,830 was 70%.

	Radioiodine was administered as: 3.7GBq and 3.7GBq in two arms			
Data sources				
Health outcomes: Quality of life was assessed using the SF-36 with acute recall period immediately before ^{131}I administration, at 6 weeks after radioiodine administration and at 3- and 8-month visits. QALYs were assessed using the EQ-5D at randomization, immediately before ^{131}I administration, 2,4, and 6 weeks after radioiodine administration and at 3- and 8-month visits. **Quality-of-life weights:** EQ-5D values collected as part of the current study were weighted using the French EQ-5D tariff. **Cost sources:** Hospital fixed, and variable costs were obtained from the French National Cost Survey using the diagnosis-related group code for ^{131}I administration. The price of rhTSH was obtained from the French drug database. Indirect costs were evaluated based on the loss of productivity incurred by sick leave using the friction cost approach. One day off work translated into 0.8 days of lost productivity to adapt the adjustment time period to absenteeism. The value of lost productivity was based on national values. Transportation costs were estimated using the French health insurance reimbursement tariffs, according to the home-hospital distance and type of transportation used. Mean cost per patient was calculated with and without indirect costs, according to the French guidelines for cost-effectiveness studies.				
Comments				
Source of funding: French Ministry of Health through the National Institute of Cancer **Limitations:** French healthcare system perspective. Discounting was not applied and not applicable given short time horizon. Utility values used to calculate QALYs were derived from EQ-5D scores using French tariff. Incremental QALY gain reported (0.013) differs from that calculated from reported total mean values for each intervention (0.012). Limited sensitivity analyses were conducted. Disclosures provided by authors were not identified online. **Other:** None				
Overall applicability:[c] Partially applicable **Overall quality:**[d] Minor limitations				

Abbreviations: 95% CI= 95% confidence interval. CUA= cost–utility analysis. da= deterministic analysis. EQ-5D= Euroqol 5 dimensions (scale: 0.0 [death] to 1.0 [full health], negative values mean worse than death). ICER= incremental cost-effectiveness ratio. NA = not applicable. NR= not reported. pa= probabilistic analysis. QALYs= quality-adjusted life years. rhTSH = recombinant human thyroid stimulating hormone. T3 = triiodothyronine. T4 = thyroxine. TSH = thyroid stimulating hormone. THW = thyroid hormone withdrawal.

(a) For studies where the time horizon is longer than the treatment duration, an assumption needs to be made about the continuation of the study effect. For example, does a difference in utility between groups during treatment continue beyond the end of treatment and if so for how long.

(b) Converted using 2013/2014 purchasing power parities[32]

(c) Directly applicable / Partially applicable / Not applicable

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

Study	**Mernagh 2010[28]**			
Study details	**Population & interventions**	**Costs**	**Health outcomes**	**Cost effectiveness**
Economic analysis: CUA (health outcome: QALYs)	**Population:** Adults who underwent total thyroidectomy for	**Total costs (mean per patient):** Intervention 1: £3,202	**QALYs (mean per patient):** Intervention 1: 0.2232	**ICER (Intervention 2 versus Intervention 1):** £890 per QALY gained (da)

Study design: Markov model adapted from Mernagh 2006 [27] **Approach to analysis:** Four health states ('pre-ablation', 'ablation', 'post-ablation', 'well') were modelled with a 1-week cycle length. **Perspective:** Canadian societal perspective with Ontario as the reference province **Time horizon:** 17 weeks **Treatment effect duration:**[(a)] **Discounting:** Costs: NA Outcomes: NA	low-risk differentiated thyroid cancer and were receiving TSH stimulation in preparation for post-thyroidectomy radioiodine ablation **Cohort settings:** Start age: NR Male: NR **Intervention 1:** Endogenous stimulation of TSH with THW prior to radioiodine ablation **Intervention 2:** Exogenous stimulation of TSH with rhTSH prior to radioiodine ablation	Intervention 2: £3,151 Incremental (2−1): £51 (95% CI: NR. p=NR) **Currency & cost year:** 2007 Canadian dollars (presented here as 2007 UK pounds[(b)]) **Cost components incorporated:** Intervention cost (2 ampoules of Thyrogen®), ablative dose of ^{131}I radioiodine, whole body scan using radioiodine, inpatient hospitalization days for patients receiving radioiodine ablation, initial and follow-up specialist visits (radiation oncologist), initial and follow-up general practitioner visits, laboratory tests (serum thyroglobulin count, thyroglobulin antibody test), daily T4 medication.	Intervention 2: 0.2808 Incremental (2−1): 0.0576 (95% CI: NR. p=NR)	95% CI: NR Probability that Intervention 2 was cost effective (£20K/30K threshold): NR **Analysis of uncertainty:** Several sensitivity analyses were conducted. However, they included societal costs and therefore it was not possible to interpret these findings from the perspective of the healthcare system.
Data sources				
Health outcomes: Equal efficacy of ablation (100%) was assumed for both interventions based on the pivotal RCT by Pacini 2006 [33]. A survey of clinicians practicing in Canada was conducted to obtain the time spent in radio-protective conditions and time interval between thyroidectomy and ablation for endogenously stimulated patients in Canada. Several resource use estimates were also obtained from this survey, including the decision to omit T3 medication during the pre-ablation period. Long-term cancer recurrence was not included based on studies which found no difference between intervention arms. Estimates regarding productivity loss were based on the earlier model by Mernagh 2006 [27] and Pacini 2006[33]. **Quality-of-life weights:** Pre-ablation utility values were obtained from 4-week SF-36 data reported by Pacini 2006[33] and transformed into SF-6D utility weights using the method described by Brazier et al. 1998. Ablation utility values were based on an assumption that this health state was 0.1 better than the pre-ablation utility				

weight. 0-4 weeks post-ablation utility values were based on 1-month SF-36 data from a pivotal RCT (data reported to be on file) transformed into SF-6D utility weights. 4-8 week post-ablation utility values were based on an assumed average of 0-4 week 'post-ablation' and 'well' health states. Well utility values were based on an assumption that patients in this state were in perfect health. **Cost sources:** Intervention cost (2 ampoules of Thyrogen) and daily T4 costs (100 µg) were obtained from the Ontario Drug Benefit Formulary 2007. The cost of an ablative dose of radioiodine was obtained from the Ontario Case Costing Initiative 2007. The cost of a whole-body scan using radioiodine and inpatient hospital day costs were obtained from the London Health Science Centre 2007. Specialist, general practitioner, and laboratory test costs were obtained from the Ontario Health Insurance Policy Schedule of Benefits 2008.
Comments
Source of funding: Genzyme Corporation **Limitations:** Canadian healthcare context. Disaggregated direct and societal results were reported for the base case but not sensitivity analyses. Utility weights estimated using SF-6D mapping algorithm. No intervention effect was applied based on results of equivalence study by Pacini 2006[33]. Ontario was used as the reference province for resource use and unit costs. Quality of life differences were estimated exclusively using Pacini 2006 trial[33] which was found to be an outlier in the clinical review as it estimated a much larger QoL loss than the other two trials available. Furthermore, Pacini 2006[33] collected QoL only twice throughout the trial, forcing the authors to heavily rely on several assumptions to model QoL changes over time. **Other:** None
Overall applicability:[(c)] Partially applicable **Overall quality:**[(d)] Potentially serious limitations

Abbreviations: 95% CI= 95% confidence interval. CUA= cost–utility analysis. da= deterministic analysis. EQ-5D= Euroqol 5 dimensions (scale: 0.0 [death] to 1.0 [full health], negative values mean worse than death). ICER= incremental cost-effectiveness ratio. NA = not applicable. NR= not reported. QALYs= quality-adjusted life years. QoL= quality of life. rhTSH = recombinant human thyroid stimulating hormone. T3 = triiodothyronine. T4 = thyroxine. TSH = thyroid stimulating hormone. THW = thyroid hormone withdrawal.

(a) For studies where the time horizon is longer than the treatment duration, an assumption needs to be made about the continuation of the study effect. For example, does a difference in utility between groups during treatment continue beyond the end of treatment and if so for how long.

(b) Converted using 2007/2008 purchasing power parities[32]

(c) Directly applicable / Partially applicable / Not applicable

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

Study	**Sohn 2015 [38]**			
Study details	**Population & interventions**	**Costs**	**Health outcomes**	**Cost effectiveness**
Economic analysis: CUA (health outcome: QALYs) **Study design:** Markov model based on Mernagh 2010 [28]	**Population:** Adults who underwent total thyroidectomy for low-risk differentiated thyroid cancer and were receiving TSH stimulation in preparation for post-	**Total costs (mean per patient):** Intervention 1: £1,031 Intervention 2: £1,800 Incremental (2−1): £769 (95% CI: NR. p=NR)	**QALYs (mean per patient):** Intervention 1: 0.245 Intervention 2: 0.281 Incremental (2−1): 0.036 (95% CI: NR. p=NR)	**ICER (Intervention 2 versus Intervention 1):** £21,357 per QALY gained 95% CI: NR Probability that Intervention 2 was cost effective (£20K/30K threshold): NR

Approach to analysis: Five health states ('pre-ablation', 'ablation', 'initial post-ablation', 'second post-ablation', and 'well') were modelled with a 1-week cycle length. **Perspective:** South Korean healthcare system **Time horizon:** 17 weeks **Treatment effect duration:**[(a)] **Discounting:** Costs: NA Outcomes: NA	thyroidectomy radioiodine ablation **Cohort settings:** Start age: NR Male: NR **Intervention 1:** Endogenous stimulation of TSH with THW prior to radioiodine ablation **Intervention 2:** Exogenous stimulation of TSH with rhTSH prior to radioiodine ablation	**Currency & cost year:** 2013 South Korean won (presented here as 2013 UK pounds[(b)]) **Cost components incorporated:** Intervention cost (2-vial kit of Thyrogen), ablative dose of radioiodine, whole body scan using radioiodine, inpatient hospitalization days for patients receiving radioiodine ablation, specialist visit (radiation oncologist), practice nurse visit, laboratory tests (TSH quantification test, serum thyroglobulin count, thyroglobulin antibody test), weekly T4 and T3 medication.		**Analysis of uncertainty:** Excluding indirect costs (i.e. loss of productivity) resulted in an incremental cost of £18,848 per QALY gained. Assuming no difference between treatment arms in hospital length of stay resulted in an incremental cost of £27,127 per QALY gained. Increasing the duration of pre-ablation health state in the rhTSH arm from 1 week to 2 weeks resulted in an incremental cost of £26,064 per QALY gained. Reducing the incremental utility difference in the pre-ablation health state by 50% (i.e. increasing the utility of the THW arm from 0.548 to 0.631) resulted in an incremental cost of £26,954 per QALY gained.
Data sources				
Health outcomes: Equal efficacy of ablation (100%) was assumed for both interventions based on the pivotal clinical trial by Pacini 2006 [33]. Patients exogenously stimulated with rhTSH were ablated 1 week following thyroidectomy. Patients endogenously stimulated were ablated at various timepoints. Patients were released from the radioprotective ward earlier when prepared for ablation with exogenous stimulation based on findings by Borget 2008 and Pacini 2006 [33]. **Quality-of-life weights:** Pre-ablation utilities were obtained from 4-week SF-36 data reported in Pacini 2006 [33] transformed into utility weights using the SF-6D method described by Brazier 1998[8] Ablation utility values were based on an assumption that this health state was 0.1 better than pre-ablation. 0-4 weeks post-ablation utility values were based on 1-month SF-36 data from pivotal RCT (data reported to be on file) transformed into SF-6D values. 4-8 week post-ablation utility values were based on an assumed average of 0-4 week post-ablation and well health states. Well health state values were based on an assumption by the Medical Services Advisory Committee. Follow-up scan utility values were based on SF-36 data reported in Schroeder 2006 transformed to SF-6D values. Thyroidectomy utility values were based on an assumption with no further detail provided. Secondary colorectal cancer utility values were based on a systematic review of available utility data by Ness 1999 Utility values for secondary bone/soft tissue and salivary gland cancer were assumed to be the same as for colorectal cancer. **Cost sources:** Intervention cost (2 ampoules of Thyrogen) was obtained from Genzyme Corporation. All other costs were obtained from the Korean Health Insurance Review Agency. The cost of weekly T4 and T3 medications were based on 125µg and 60µg daily doses, respectively.				

Comments
Source of funding: Genzyme Corporation **Limitations:** Korean healthcare context. Discounting was not applied and not applicable given 17-week time horizon. Utility weights estimated using SF-6D mapping algorithm. No intervention effect was applied based on results of equivalence study by Pacini 2006. Cost year not reported and assumed to be 2013 based on unit cost reference dates. Quality of life differences were estimated exclusively using Pacini 2006 trial[33] which was found to be an outlier in the clinical review as it estimated a much larger QoL loss than the other two trials available. Furthermore, Pacini 2006[33] collected QoL only twice throughout the trial, forcing the authors to heavily rely on several assumptions to model QoL changes over time. Conflict of interest declaration was unclear - the supervising author is a medical advisor in Genzyme Corporation which funded the study. **Other:** None.
Overall applicability:(c) Partially applicable **Overall quality:**(d) Potentially serious limitations

Abbreviations: 95% CI= 95% confidence interval. CUA= cost–utility analysis. da= deterministic analysis. EQ-5D= Euroqol 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. pa= probabilistic analysis. QALYs= quality-adjusted life years. rhTSH = recombinant human thyroid stimulating hormone. T3 = triiodothyronine. T4 = thyroxine. THW = thyroid hormone withdrawal.

(a) For studies where the time horizon is longer than the treatment duration, an assumption needs to be made about the continuation of the study effect. For example, does a difference in utility between groups during treatment continue beyond the end of treatment and if so for how long.

(b) Converted using 2013/2014 purchasing power parities[32]

(c) Directly applicable / Partially applicable / Not applicable

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

Study	Vallejo 2017 [43]			
Study details	**Population & interventions**	**Costs**	**Health outcomes**	**Cost effectiveness**
Economic analysis: CUA (health outcome: QALYs) **Study design:** Markov model based on Mernagh 2010 [28] **Approach to analysis:** Five health states ('pre-ablation', 'ablation', 'initial post-ablation', 'second post-ablation', and 'well') were	**Population:** Adults who underwent total thyroidectomy for low-risk differentiated thyroid cancer and were receiving TSH stimulation in preparation for post-thyroidectomy radioiodine ablation **Cohort settings:** Start age: NR Male: NR	**Total costs (mean per patient):** Intervention 1: £5,337 Intervention 2: ££4,697 Incremental (2−1): -£640 (95% CI: NR. p=NR) **Currency & cost year:** 2015 Spanish euros (presented here as 2015 UK pounds(b))	**QALYs (mean per patient):** Intervention 1: 0.233 Intervention 2: 0.281 Incremental (2−1): 0.048 (95% CI: NR. p=NR)	**ICER (Intervention 2 versus Intervention 1):** Domainates (greater QALY gain at a lower cost) 95% CI: NR Probability that Intervention 2 was cost effective (£20K/30K threshold): NR **Analysis of uncertainty:** Assuming no difference between treatment arms in hospital length of stay resulted in an incremental cost of £1,057 per QALY gained.

modelled with a 1-week cycle length. **Perspective:** Spanish healthcare system **Time horizon:** 17 weeks **Treatment effect duration:**[(a)] **Discounting:** Costs: NA Outcomes: NA	**Intervention 1:** Endogenous stimulation of TSH with THW prior to radioiodine ablation **Intervention 2:** Exogenous stimulation of TSH with rhTSH prior to radioiodine ablation	**Cost components incorporated:** Intervention cost (2-vial kit of Thyrogen), ablative dose of radioiodine, whole body scan using radioiodine, inpatient hospitalization days for patients receiving radioiodine ablation, specialist visit (radiation oncologist), practice nurse visit, laboratory tests (TSH quantification test, serum thyroglobulin count, thyroglobulin antibody test), weekly T4 and T3 medication.		
Data sources				
Health outcomes: Equal efficacy of ablation (100%) was assumed for both interventions based on the pivotal RCT by Pacini 2006 [33]. Time in each health state was obtained from a survey of medical specialists at 20 public and private Spanish healthcare centres conducted as part of the current study. Duration of stay in metabolic therapy room was obtained from a study by Borget 2008. **Quality-of-life weights:** Pre-ablation utility values were obtained from 4-week SF-36 data reported by Pacini 2006 and transformed into SF-6D utility weights using the method described by Brazier et al. 1998[8]. Quality of life differences were estimated exclusively using Pacini 2006 trial[33] which was found to be an outlier in the clinical review as it estimated a much larger QoL loss than the other two trials available. Furthermore, Pacini 2006[33] collected QoL only twice throughout the trial, forcing the authors to heavily rely on several assumptions to model QoL changes over time. Initial post-ablation utility values were based on 1-month SF-36 data from a pivotal RCT (data reported to be on file) transformed into SF-6D utility weights. Secondary post-ablation utility values were based on an assumed average of 'initial post-ablation' and 'well' health states. Well utility values were based on an assumption that patients in this state were in perfect health. **Cost sources:** NR.				
Comments				
Source of funding: Sanofi-Genzyme **Limitations:** Spanish healthcare context. Discounting was not applied and not applicable given 17-week time horizon. Utility weights estimated using SF-6D mapping algorithm. No intervention effect was applied based on results of equivalence study by Pacini 2006 **Other:** None.				
Overall applicability:[(c)] Partially applicable **Overall quality:**[(d)] Potentially serious limitations				

Abbreviations: 95% CI= 95% confidence interval. CUA= cost–utility analysis. da= deterministic analysis. EQ-5D= Euroqol 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. pa= probabilistic analysis. QALYs= quality-adjusted life years. QoL= quality of life. rhTSH = recombinant human thyroid stimulating hormone. T3 = triiodothyronine. T4 = thyroxine. THW = thyroid hormone withdrawal.

(a) For studies where the time horizon is longer than the treatment duration, an assumption needs to be made about the continuation of the study effect. For example, does a difference in utility between groups during treatment continue beyond the end of treatment and if so for how long.
(b) Converted using 2015/2016 purchasing power parities[32]
(c) Directly applicable / Partially applicable / Not applicable
(d) Minor limitations / Potentially serious limitations / Very serious limitations

Appendix I – Excluded studies

I.1 Clinical studies

Table 11: Studies excluded from the clinical review

Reference	Reason for exclusion
Barbaro 2007[2]	Inappropriate study design – systematic review
Barbaro 2003[3]	Inappropriate study design – cohort study
Barbaro 2006[4]	Inappropriate study design – cohort study
Campenni 2018[9]	Inappropriate study design – systematic review
Doi 2000[13]	Inappropriate study design – systematic review
Iakovou 2016[18]	Inappropriate study design – cohort study
Lamartina 2015[20]	Inappropriate study design – systematic review
Lizuka, 2020[19]	Inappropriate study design – cohort study
Ma 2010[23]	Inappropriate study design – systematic review
Mallick 2008[24]	Inappropriate study design – review article
Marturano 2015[26]	Inappropriate intervention – 131I scan only
Mernagh 2006[27]	Inappropriate study design – health economics study
Mernagh 2010[28]	Inappropriate study design – health economics study
Nygaard 2013[31]	Inappropriate study design – cross-over study
Pacini 2002[34]	Inappropriate study population – mixed population (ablation and non-ablation patients)
Pak 2014[35]	Inappropriate study design – systematic review
Robbins 2002[36]	Inappropriate study design – cohort study
Sohn 2015[38]	Inappropriate study design – health economics study
Taieb 2010[40]	Inappropriate comparison – no relevant outcomes
Tu 2014[41]	Inappropriate study design – systematic review
Vaiano 2007[42]	Inappropriate comparison – no relevant outcomes
van der Horst-Schrivers 2015[44]	Inappropriate study design – cohort study
Verburg 2020[45]	Inappropriate study design – systematic review
Xu 2015[47]	Inappropriate study design – systematic review
Yoo 2009[48]	Inappropriate study design – systematic review

I.2 Health Economic studies

Table 12: Studies excluded from the health economic review

Reference	Reason for exclusion
Blamey 2005[5]	Excluded as rated not applicable. The population was people receiving rhTSH for diagnostic purposes, not in preparation for RAI.
Mernagh 2006[27]	Excluded as rated not applicable. Total or incremental costs could not be extracted for an NHS perspective only and indirect costs accounted for the majority of the total costs. In addition, a more applicable analysis[28] was available based on the same RCT this study was selectively excluded.

Reference	Reason for exclusion
Waissi 2019[46]	Excluded as rated not applicable. Total costs were from a societal perspective only.

1