THYROID
Volume 35, Number 10, 2025
© Mary Ann Liebert, Inc.
DOI: 10.1177/10507256251372193

REVIEWS and SCHOLARLY DIALOG

Mortality Risks Associated with Antithyroid Drugs, Radioactive Iodine, and Surgery for Hyperthyroidism: A Systematic Review and Network Meta-Analysis

Carol Chiung-Hui Peng,[1] Brianna R. Spiegel,[2] David Flynn,[3] Huei-Kai Huang,[4–6] Ching-Hui Loh,[6,7] Peter Pin-Sung Liu,[8,*] and Elizabeth N. Pearce[9,*]

Introduction: Hyperthyroidism can be treated with antithyroid drugs (ATD), radioactive iodine (RAI), or surgery. We aimed to evaluate the long-term outcomes of these treatments through a systematic review and network meta-analysis (NMA).
Methods: A systematic literature search of PubMed, EMBASE, Web of Science, and the Cochrane Library (from inception to March 7, 2025) was conducted to identify studies comparing the risks of all-cause mortality, cardiovascular mortality, major adverse cardiovascular events (MACE), and cancer mortality among patients with hyperthyroidism treated with ATD, RAI, or surgery. Pooled effect estimates were expressed as hazard ratios (HR) with confidence intervals (CI) using a random-effects model. The study was registered with PROSPERO (CRD420250543380) and adhered to the PRISMA-NMA guidelines.
Results: Of the 8163 studies screened, 12 observational studies with an overall moderate risk of bias, comprising 192,208 patients were included in this NMA. Most patients received ATD ($n = 142{,}622$), followed by RAI ($n = 19{,}303$) and surgery ($n = 10{,}360$). Surgery was associated with decreased risks of all-cause mortality and cardiovascular mortality compared with both ATD and RAI. For all-cause mortality, the pooled HRs (CI, p-values) were 0.58 (0.45–0.75, $p < 0.0001$) for surgery versus ATD and 0.68 (0.56–0.84, $p = 0.0004$) for surgery versus RAI. For cardiovascular mortality, the pooled HR (CI, p-values) were 0.43 (0.19–0.98, $p = 0.0445$) for surgery versus ATD and 0.55 (0.33–0.93, $p = 0.0269$) for surgery versus RAI. No significant differences were observed in MACE or cancer mortality across the treatment groups.
Conclusions: In patients with hyperthyroidism, surgery was associated with significantly decreased risks of all-cause mortality and cardiovascular mortality compared with ATD and RAI. Risks of MACE and cancer mortality did not differ by type of hyperthyroidism treatment. However, these findings should be interpreted with caution due to inherent methodological limitations of observational studies, including, but not limited to heterogeneity and potential selection bias.

Keywords: hyperthyroidism, mortality, antithyroid drugs, radioactive iodine, thyroidectomy, network meta-analysis

[1]Division of Endocrinology and Metabolism, Department of Internal Medicine, Chi Mei Medical Center, Tainan, Taiwan.
[2]Boston University Chobanian and Avedisian School of Medicine, Boston, Massachusetts, USA.
[3]Department of Medical Sciences & Education, Boston University Chobanian and Avedisian School of Medicine, Boston, Massachusetts, USA.
[4]Department of Family Medicine, Hualien Tzu Chi Hospital, Buddhist Tzu Chi Medical Foundation, Hualien, Taiwan.
[5]Center for Clinical Epidemiology and Biostatistics, Hualien Tzu Chi Hospital, Buddhist Tzu Chi Medical Foundation, Hualien, Taiwan.
[6]School of Medicine, Tzu Chi University, Hualien, Taiwan.
[7]Center for Aging and Health, Hualien Tzu Chi Hospital, Buddhist Tzu Chi Medical Foundation, Hualien, Taiwan.
[8]Population Health Data Center, National Cheng Kung University, Tainan, Taiwan.
[9]Section of Endocrinology, Diabetes, Nutrition and Weight Management, Boston University Chobanian and Avedisian School of Medicine, Boston, Massachusetts, USA.
*Contributed equally.

Introduction

Hyperthyroidism, a state of excessive thyroid hormone synthesis, is usually due to Graves' disease or toxic nodular goiter.[1,2] Its global prevalence ranges from 0.2% to 2.5% in iodine-sufficient areas to 10–15% in iodine-deficient areas.[1,2] Hyperthyroidism is linked to higher all-cause and cardiovascular mortality risks.[3–5] These risks increase with prolonged uncontrolled hyperthyroidism, even in treated patients.[6,7]

Hyperthyroidism treatment options include antithyroid drugs (ATD), radioactive iodine (RAI), or surgery. A recent survey across 85 countries highlighted regional variations and a paradigm shift in managing Graves' disease.[8] Most clinicians (91.5%) preferred ATD, followed by RAI (7%) and surgery (1.5%).[8] In the United States, RAI use as the initial therapy has decreased from 69% to 11% over the past three decades,[8] possibly reflecting clinician concerns about thyroid eye disease and patients' preferences, driven by fear of radioactivity, risk of cancer and pregnancy-related concerns, often independent of actual risk.[9–12] Although American Thyroid Association (ATA) guidelines for hyperthyroidism recommend considering RAI or surgery if remission is not achieved with ATD after 12–18 months,[13] most surveyed clinicians favored prolonged ATD use.[8,14,15] Most providers considered surgery after ATD only for patients with thyroid eye disease or planning conception.[8]

This shift in practice is driven more by the desire to control hyperthyroidism without causing iatrogenic hypothyroidism than by strong evidence regarding long-term risks of each approach.[3] While emerging data support the safety of long-term ATD use,[7,8] two recent large-scale population-based cohorts suggested that surgery and RAI are associated with lower all-cause mortality, cardiovascular diseases, and major adverse cardiovascular events (MACE) risks compared with ATD.[16,17] No comprehensive analysis directly compares the long-term outcomes of these three treatments. To address this gap, we conducted a network meta-analysis to evaluate all-cause mortality, cardiovascular mortality, MACE, and cancer mortality in hyperthyroid patients treated with ATD, RAI, or surgery.

Materials and Methods

Protocol and data sources

This network meta-analysis was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Network Meta-Analyses (PRISMA-NMA) guidelines.[18] A medical librarian performed a comprehensive literature search across four electronic databases—PubMed, EMBASE, Web of Science, and the Cochrane Library—from inception to March 7, 2025. Search strategies were adapted for each database using predefined terms related to hyperthyroidism, goiter, ATD, thyroidectomy, RAI, and mortality. Both keyword-based and subject-heading approaches were applied. No publication date restrictions were placed. Studies were limited to those conducted in humans and published in English. The full search strategy is provided in Supplementary Table S1. We manually checked reference lists of relevant reviews and meta-analyses for additional articles. This systematic review and NMA was registered in PROSPERO (CRD420250543380).

Study selection and outcomes

Eligibility criteria were established using the Population, Intervention, Comparator, and Outcome (PICO) framework: (1) Population: Patients with hyperthyroidism. (2) Interventions and comparators: ATD, RAI, or surgery (thyroidectomy). (3) Primary outcome: All-cause mortality; secondary outcomes: cardiovascular mortality, MACE, and cancer mortality. We accepted the definitions of MACE provided in each study. Eligible studies included clinical trials, cohort, or case-control studies published as original articles with original data. Studies had to compare at least two hyperthyroidism treatments and report relevant outcomes. We excluded studies evaluating a single treatment, comparing treated patients with untreated individuals or healthy controls, or lacking sufficient risk estimate data. Laboratory or animal studies, review articles, case reports, editorials, letters, and conference abstracts without original findings or adequate data, were excluded.

The systematic review was conducted using Covidence (Covidence systematic review software, Veritas Health Innovation, Melbourne, Australia. Available at www.covidence.org). Two independent reviewers (B.R.S. and C.C.P.) screened retrieved articles based on titles and abstracts. If either reviewer deemed a study potentially eligible, it proceeded to full-text review. Full-text assessments were independently conducted by the same two reviewers, with discrepancies resolved through discussion with a third reviewer (P.P.L.).

Data extraction and quality assessment

Two reviewers (B.R.S. and C.C.P.) independently performed data extraction using a predefined standardized form, collecting information on (1) study characteristics: first author, publication year, study design, geographic location, data source, observation period; (2) patient characteristics: sample size, mean age, sex distribution, hyperthyroidism etiology, and treatment strategies; (3) outcomes: all-cause mortality, cardiovascular mortality, MACE, and cancer mortality, along with risk estimates and their corresponding confidence intervals; (4) statistical adjustments: confounders and adjustment methods. A third reviewer (P.P.L.) verified data accuracy.

Study quality was independently assessed by two reviewers (B.R.S. and P.P.L.) using the Newcastle–Ottawa Scale (NOS) for observational studies, evaluating selection bias, comparability, and outcome assessment. Discrepancies were resolved by consensus with a third reviewer (C.C.P.).

Data synthesis and statistical analysis

The hazard ratio (HR) with confidence intervals (CI) was extracted as the effect measure. If multiple risk estimates were reported (e.g., covariate-adjusted vs. propensity score-adjusted models), propensity score matching or weighting estimates were prioritized. A random-effects network meta-analysis was conducted using a frequentist approach to estimate pooled HR. Since all included studies were observational and individual patient-level data were not available, we employed the frequentist framework, which is suitable for addressing between-study heterogeneity without relying on strong prior assumptions typically required by Bayesian analysis. Treatment ranking was determined using P-scores, which estimate the probability that one treatment is superior to another across all comparisons. A higher P-score indicates a lower likelihood of mortality or other adverse outcomes associated with that treatment. The Greenland and Longnecker method was employed to handle the balance of multi-arm results to fulfill the transitivity assumption required for network meta-analysis.[19]

Heterogeneity was assessed using the I^2, τ^2, and Q statistic, categorized as low ($I^2 < 50\%$), moderate ($50\% \leq I^2 \leq 75\%$), or high ($I^2 > 75\%$). Network consistency was evaluated using node-splitting analysis, comparing direct and indirect estimates. When a difference is $p < 0.05$, there is a significant inconsistency between the direct and indirect estimates. Publication bias was assessed via funnel plots with Egger's test.

Sensitivity analyses were performed to test robustness by excluding high-risk studies and restricting analyses to studies with confounder-adjusted results. We used the leave-one-out method to assess whether the overall effect estimate changed when one study was removed at a time.

Statistical analyses were conducted using R version 4.4.0 (R Core Team, Vienna, Austria) with the netmeta package. Two-sided *p*-values <0.05 were considered statistically significant.

Results

We identified 6696 references after removing duplicates, with 94 full-text articles reviewed. Ron et al. (1998) and Kitahara et al. (2020) both reported long-term cancer mortality outcomes from the Cooperative Thyrotoxicosis Therapy Follow-up Study (CTTFUS).[20,21] To avoid duplication, only Kitahara et al.'s study was included, as it provided longer follow-up of the same cohort. Goldman et al. also reported results from a small portion of the same cohort; their study was retained for analyzing all-cause mortality but excluded from cancer mortality analysis.[22] Ultimately, 12 eligible studies were included in the current network-meta analysis.[7,16,17,21–29] The PRISMA flow diagram is shown in Figure 1. Included studies spanned seven countries across Europe, North America, Asia, and the Middle East. Sample sizes ranged from 963 to 114,062 with follow-up durations of 4.4 to 26.0 years. Nine studies reported all-cause mortality, three reported cardiovascular mortality, two reported MACE, and four reported cancer mortality. Study characteristics, including age, sex, sample size, and study design, are summarized in Table 1. No randomized controlled trial (RCT) comparing three modalities was identified in our systematic review. Adjusted covariates for each study are provided in Supplementary Table S2; all of the included studies controlled for patient characteristics, 11 controlled for characteristics of hyperthyroidism, 7 controlled for comorbidities, and 4 controlled for coexisting medications. Quality assessments using NOS are in Supplementary Table S3.

All-cause mortality

The network structure for each outcome is shown in Figure 2. The numbers of studies comparing all-cause mortality between treatments were: RAI versus ATD ($n = 6$), surgery versus ATD ($n = 3$), and surgery versus RAI ($n = 6$). Most patients received ATD ($n = 143{,}622$), followed by RAI ($n = 19{,}303$) and surgery ($n = 10{,}630$). Figure 3 presents results from individual studies.

The pooled HRs (CI, *p* value) for all-cause mortality were 0.58 (0.45–0.75, <0.0001) when comparing surgery versus ATD, 0.85 (0.70–1.03, 0.0893) when comparing RAI versus ATD, and 0.68 (0.56–0.84, 0.0004) when comparing surgery versus RAI (Fig. 4). Overall heterogeneity was moderate ($I^2 = 73.1\%$, *p* value <0.0001). Surgery had the highest P-score (0.9999) for all-cause mortality (Table 2). Node-splitting plots (Supplementary Fig. S1) showed no significant inconsistency. No significant publication bias was detected (Egger test, $p = 0.0781$; Supplementary Fig. S2).

Cardiovascular mortality, MACE, and cancer mortality outcomes

The pooled HRs (CI, *p* value) for cardiovascular mortality were 0.43 (0.19–0.98, 0.0445) when comparing surgery versus ATD, 0.77 (0.32–1.86, 0.5624) when comparing RAI versus ATD, and 0.55 (0.33–0.93, 0.0269) when comparing

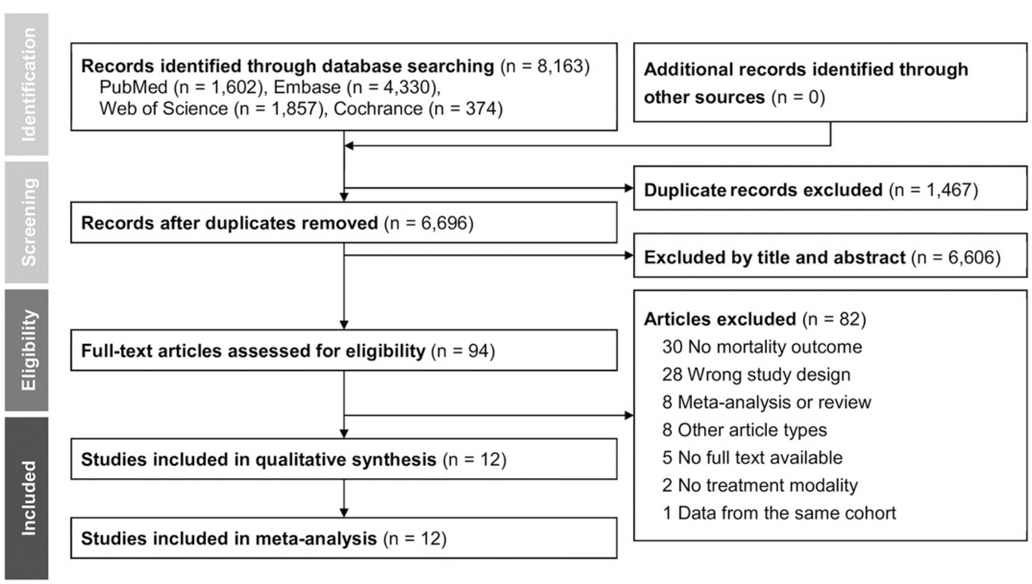

FIG. 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram.

Table 1. Characteristics of Included Studies

Author (year)	Country	Data source	Etiology of hyperthyroidism	Sample size	Mean age (years)	Female (%)	Patient numbers: ATD	Patient numbers: RAI	Patient numbers: Surgery	Mean age at treatment: ATD	Mean age at treatment: RAI	Mean age at treatment: Surgery	Subgroup of post RAI thyroid function	Type of thyroidectomy	Mean follow-up time (years)	Outcome
Boelaert et al. (2013)[23]	UK	Single-center	Graves' disease (+TRAb) Toxic nodule Others	1036	57.0	78.8	272	764	—	54.0	59.0	—	Hypothyroid vs. not hypothyroid (followed for 5 years after diagnosis)	—	12.4	All-cause mortality
Giesecke et al. (2018)[24]	Sweden	National database	Graves' disease Toxic nodule	10,992	62.9	84.6	—	10,250	742	—	64.0	47.0	NA	NA	16.3	All-cause mortality Cancer mortality CV mortality
Goldman et al. (1988)[22]	USA	Single-center	Not specified	963	45.2	100.0	—	607	356	—	50.6	36.1	NA	NA	17.2	All-cause mortality Cancer mortality
Gronich et al. (2020)[25]	Israel	Health service database	Thyrotoxicosis with or without goiter (ICD-9: 242) Thyroid nodular goiter (ICD-9: 241) Benign thyroid neoplasm (ICD-9: 226)	16,637	51.9	74.0	13,808	2829	—	51.7	52.8	—	NA	—	7.4	All-cause mortality
Hoffman et al. (1982)[26]	USA	Single-center	Graves' disease Toxic nodule Others	3146	68.6	100.0	—	1,005	2,141	—	56.8	45.7	NA	NA	19.0	All-cause mortality Cancer mortality
Kitahara et al. (2020)[21]	USA and UK	Multicenter	Graves' disease Toxic nodule Others	19,229	45.5	80.0	1,138	7474	10,617	45.6	50.8	42	NA	NA (but suggested 89.0% Subtotal)*	26.0	Cancer mortality
Liu et al. (2021)[16]	Hong Kong	Regional database	Graves' disease (ICD-9: 242.00/ 242.01)	6385	45.6	72.6	4784	1274	327	44.6	49.2	41.0	NA	82% Total 18% Subtotal	7.4	All-cause mortality
Liu et al. (2022)[27]	Hong Kong	Regional database	Graves' disease (ICD-9: 242.00/ 242.01)	3443	45.1	73.0	2294	755	394	44.6	47.9	42.3	NA	87% Total 13% Subtotal	6.5	All-cause mortality
Okosieme et al. (2019)[7]	UK	National database	Graves' disease (+TRAb)	4019	48.0	81.5	3587	432	—	48.0	50.0	—	Controlled† vs. hyperthyroid (1 years after diagnosis)	—	6.0	All-cause mortality MACE‡
Peng et al. (2024)[17]	Taiwan	National database	Graves' disease Toxic nodule Others	114,062	44.1	73.2	107,052	1238	5772	44.0	44.9	46.6	NA	Total 17.7% Subtotal 37.5% Others 44.8%	4.4	All-cause mortality CV mortality MACE§
Ryödi et al. (2015)[28]¶	Finland	Single-center and National database	Graves' disease Toxic nodule Others	6148	49.8	84.6	—	1814	4334	—	59.0	46.0	NA	Total 31% Subtotal 69%+	10.6	Cancer mortality
Ryödi et al. (2018)[29]¶	Finland	Single-center and National database	Graves' disease Toxic nodule Others	6148	49.8	84.6	—	1814	4334	—	59.0	46.0	Hypothyroid vs. not hypothyroid (started levothyroxine during follow-up)	Total 68% Subtotal 31%+	10.6	CV Mortality

None of the included studies were randomized controlled trials.

*For 2718 patients (89.0%) treated with both surgery and RAI, the date of surgery preceded the date of the first RAI treatment, suggesting these were primarily patients who underwent subtotal thyroidectomy and subsequently experienced relapse.

†Definition of controlled hyperthyroidism: Initiation of levothyroxine, thyroid-stimulating hormone level above the reference range, or persistent euthyroidism (>6 months), each occurring without concurrent antithyroid drug use.

‡Acute myocardial infarction, congestive heart failure, ischemic cerebrovascular accident, or death (no individual cardiovascular mortality result reported).

§Acute myocardial infarction, stroke, heart failure, cardiovascular mortality.

¶Ryodi 2015 and 2018: Both studies analyzed the same cohort with an identical number of thyroidectomies but reported opposite proportions of patients undergoing total versus near-total thyroidectomy.

ATD, antithyroid drugs; RAI, radioactive iodine; NA, not applicable.

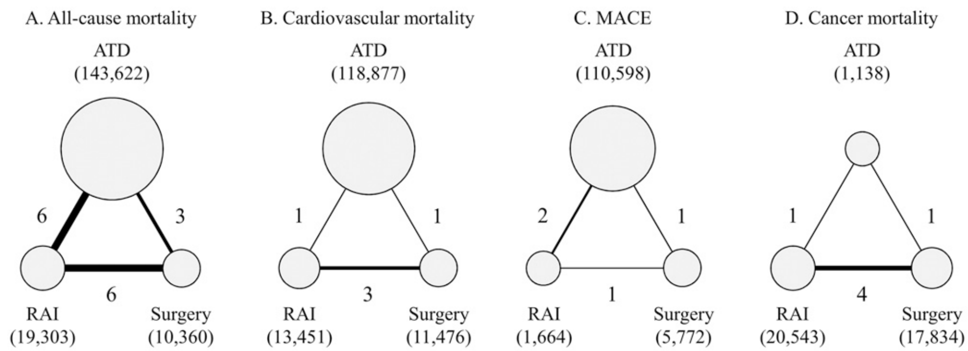

FIG. 2. Network diagram of treatment comparisons for **(A)** all-cause mortality, **(B)** cardiovascular mortality, **(C)** MACE outcomes, and **(D)** cancer mortality. The size of the nodes is proportional to the number of patients (in parentheses) who received the treatment. The width of the lines is proportional to the number of studies comparing the connected treatments. ATD, antithyroid drugs; RAI, radioactive iodine; MACE, major adverse cardiovascular events.

surgery versus RAI (Fig. 4). The overall heterogeneity was high ($I^2 = 85.6\%$, p value = 0.0010).

The pooled HRs (CI, p value) for MACE were 0.88 (0.35–2.20, 0.7842) when comparing surgery versus ATD, 0.71 (0.34–1.51, 0.3743) when comparing RAI versus ATD, and 1.24 (0.44–3.44, 0.6847) when comparing surgery versus RAI (Fig. 4). Overall heterogeneity was moderate ($I^2 = 74.5\%$, p value = 0.0478).

The pooled HRs (95% CI, p value) for cancer mortality were 0.99 (0.87–1.13, 0.9129) when comparing surgery versus ATD, 1.01 (0.88–1.16, 0.8665) when comparing RAI versus ATD, and 0.98 (0.93–1.04, 0.5189) when comparing surgery versus RAI (Fig. 4). Overall heterogeneity was low ($I^2 = 0\%$, p value = 0.6398).

No significant inconsistency was observed for cardiovascular mortality (Supplementary Fig. S3) or MACE (Supplementary Fig. S4) outcomes, while an inconsistency was detected in the RAI versus ATD comparison for cancer mortality (Supplementary Fig. S5).

Sensitivity analyses

We performed sensitivity analyses after excluding studies that only provided unadjusted estimates for all-cause mortality (Supplementary Fig. S6).[20] Only Goldman (1988)[22] was excluded from the all-cause mortality analysis because the original effect size was not adjusted for any confounders. The pooled HRs (CI, p value) for all-cause mortality were 0.66 (0.52–0.82, 0.0002) when comparing surgery versus ATD, 0.84 (0.73–0.98, 0.0285) when comparing RAI versus ATD, and 0.78 (0.64–0.94, 0.0084) when comparing surgery versus RAI. Overall heterogeneity was moderate ($I^2 = 58.2\%$, p value = 0.0104).

As a post-hoc sensitivity analysis, we excluded the large-scale study by Peng et al. (2024), which contributed 59% of the total sample size, to assess its potential influence on the findings. Results remained generally consistent with the main analysis for all-cause mortality and cancer mortality. Insufficient studies were available to perform NMA for cardiovascular mortality and MACE. (Supplementary Fig. S7) The leave-one-out meta-analysis for all-cause mortality (Supplementary Fig. S8) showed no change in the pooled effect estimate upon exclusion of any single study.

Discussion

In this systematic review and network meta-analysis of 12 studies including 192,208 participants, surgery for hyperthyroidism was associated with a 42% lower all-cause mortality risk and a 57% lower cardiovascular mortality risk compared with ATD. Compared with RAI, surgery was linked to a 32% lower risk of all-cause mortality and a 45% lower risk of cardiovascular mortality. However, no significant differences were observed in the risk of MACE or cancer mortality among patients treated with surgery, RAI, or ATD.

Despite emerging evidence supporting better long-term outcomes with surgery compared with RAI and ATD,[16,17] thyroidectomy remains the last resort for treating hyperthyroidism, likely due to concerns about scarring, cost, limited access to high-volume surgeons, and risks of serious complications such as recurrent laryngeal nerve injury and hypoparathyroidism.[30] Surgery is used in 2%–6% of cases globally, often reserved for patients with thyroid eye disease or planning pregnancy.[8,31,32] Our analysis showed that surgery was associated with lower risks of all-cause and cardiovascular mortality compared to both ATD and RAI. While selection bias may have contributed to these findings, as surgical patients were often younger than those receiving RAI or ATD, this was not the case in Peng et al., where the surgical cohort was slightly older. Moreover, all studies except Goldman et al. adjusted for age, addressing concerns that age differences might confound mortality outcomes. Sensitivity analysis excluding Goldman et al. confirmed the robustness of our results.

A longer duration without achieving euthyroidism, or without inducing and treating hypothyroidism, has been associated with increased mortality, even among treated patients.[6] Therefore, both the type and timing of surgery during hyperthyroidism treatment may influence long-term outcomes. Although ATA guidelines recommend total thyroidectomy for Graves' disease to minimize relapse risk,[13] the proportion of total thyroidectomies varied across studies, ranging from 68% to 87% in those reporting surgical types. Two notable outliers included the CTTFUS cohort, where 89% underwent subtotal thyroidectomy (reflecting earlier practices before the 1970s)·[21] and the Taiwanese cohort where only 18% received total thyroidectomy.[17] Regarding surgical timing, ATA guidelines suggest considering thyroidectomy or RAI for Graves' disease if

Outcome	Treatment vs comparator	Study	HR (95% CI)	*p* -value
All-cause mortality	Surgery vs ATD	Liu (2021)	0.39 (0.22–0.68)	0.0011
		Liu (2022)	0.42 (0.24–0.73)	0.0022
		Peng (2024)	0.54 (0.34–0.86)	0.0092
	RAI vs ATD	Peng (2024)	0.68 (0.45–1.02)	0.0647
		Boelaert (2013)	0.82 (0.66–1.02)	0.0739
		Gronich (2020)	0.83 (0.72–0.95)	0.0084
		Liu (2021)	0.94 (0.77–1.14)	0.5365
		Okosieme (2019)	0.96 (0.68–1.35)	0.8155
		Liu (2022)	1.10 (0.84–1.44)	0.4882
	Surgery vs RAI	Liu (2022)	0.38 (0.21–0.68)	0.0012
		Liu (2021)	0.41 (0.23–0.74)	0.0028
		Goldman (1988)	0.50 (0.38–0.64)	<.0001
		Peng (2024)	0.79 (0.50–1.27)	0.3216
		Giesecke (2018)	0.85 (0.72–0.99)	0.0454
		Hoffman (1982)	1.00 (0.87–1.15)	1.0000
Cardiovascular mortality	Surgery vs ATD	Peng (2024)	0.45 (0.28–0.72)	0.0009
	RAI vs ATD	Peng (2024)	0.70 (0.32–1.56)	0.3774
	Surgery vs RAI	Ryödi (2018)	0.39 (0.32–0.48)	<.0001
		Peng (2024)	0.64 (0.26–1.62)	0.3389
		Giesecke (2018)	0.74 (0.56–0.97)	0.0317
MACE	Surgery vs ATD	Peng (2024)	0.76 (0.60–0.97)	0.0251
	RAI vs ATD	Peng (2024)	0.45 (0.22–0.91)	0.0275
		Okosieme (2019)	0.98 (0.73–1.33)	0.8950
	Surgery vs RAI	Peng (2024)	1.69 (0.80–3.55)	0.1675
Cancer mortality	Surgery vs ATD	Kitahara (2020)	0.99 (0.87–1.13)	0.8802
	RAI vs ATD	Kitahara (2020)	1.03 (0.89–1.18)	0.6812
	Surgery vs RAI	Kitahara (2020)	0.97 (0.91–1.03)	0.3351
		Ryödi (2015)	1.04 (0.78–1.38)	0.7876
		Giesecke (2018)	1.07 (0.82–1.41)	0.6246
		Hoffman (1982)	1.25 (0.78–2.02)	0.3580

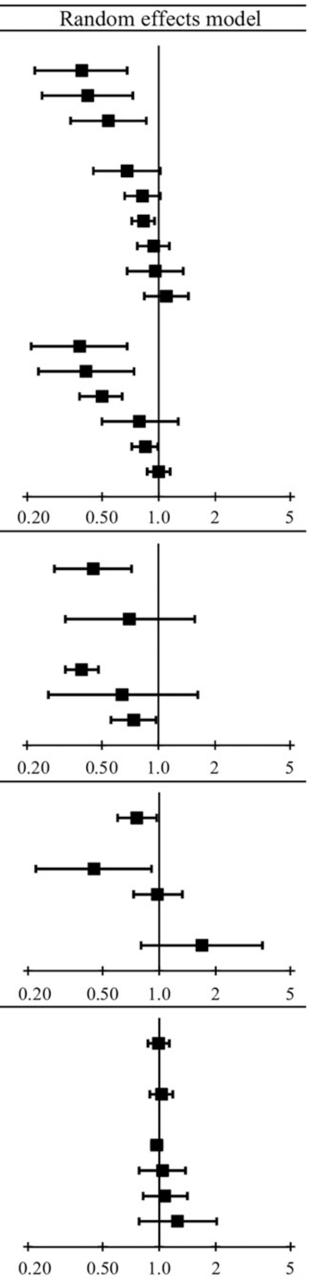

FIG. 3. Hazard ratios of included studies for all-cause mortality, cardiovascular mortality, MACE, and cancer mortality outcomes. The order of enrolled studies is sorted by the treatment effect size. ATD, antithyroid drugs; RAI, radioactive iodine; MACE, major adverse cardiovascular event; HR, hazard ratio; CI, confidence interval.

remission is not achieved after 12 to 18 months of ATD.[13] Liu et al. published two studies: one examining surgery within one year of ATD use[16] and another evaluating surgery after one year of treatment.[27] Similarly, Peng et al. analyzed outcomes based on surgery performed within 18 months of diagnosis, with a sensitivity analysis for surgery after 24 months.[17] Both studies found consistent benefits regardless of surgical timing, suggesting that hypothyroidism induction may be the key factor in improving long-term outcomes.

Although RAI was associated with higher all-cause and cardiovascular mortality risks compared with surgery, increased risks associated with hyperthyroidism are mitigated following RAI, provided hypothyroidism is appropriately managed with levothyroxine.[7,23,29,33] ATA guidelines recommend aiming for near-complete ablation and permanent hypothyroidism in Graves' disease, while for toxic nodules, the goal is to ablate autonomous tissue and preserve normal thyroid function.[13] RAI effects are not immediate. It often causes a transient rise in thyroid hormone levels, followed by a gradual return to euthyroidism or progression to hypothyroidism within 4–10 weeks, or even longer.[1,2] Persistent hyperthyroidism after RAI can result from various factors, and hypothyroidism eventually occurs in 50%–85% of patients with Graves' disease.[1,2] Consequently, patients receiving RAI may remain hyper- or hypothyroid for extended periods, unlike those undergoing surgery, who become hypothyroid immediately and can begin

Outcome	Treatment vs comparator	pooled HR (95% CI)	p-value	I^2	Q	τ^2
All-cause mortality	Surgery vs ATD	0.58 (0.45–0.75)	<.0001	73.1%	37.2	0.0410
	RAI vs ATD	0.85 (0.70–1.03)	0.0893			
	Surgery vs RAI	0.68 (0.56–0.84)	0.0004			
Cardiovascular mortality	Surgery vs ATD	0.43 (0.19–0.98)	0.0445	85.6%	13.9	0.1629
	RAI vs ATD	0.77 (0.32–1.86)	0.5624			
	Surgery vs RAI	0.55 (0.33–0.93)	0.0269			
MACE	Surgery vs ATD	0.88 (0.35–2.20)	0.7842	74.5%	3.9	0.2256
	RAI vs ATD	0.71 (0.34–1.51)	0.3743			
	Surgery vs RAI	1.24 (0.44–3.44)	0.6847			
Cancer mortality	Surgery vs ATD	0.99 (0.87–1.13)	0.9129	0.0%	1.7	0.0000
	RAI vs ATD	1.01 (0.88–1.16)	0.8665			
	Surgery vs RAI	0.98 (0.93–1.04)	0.5189			

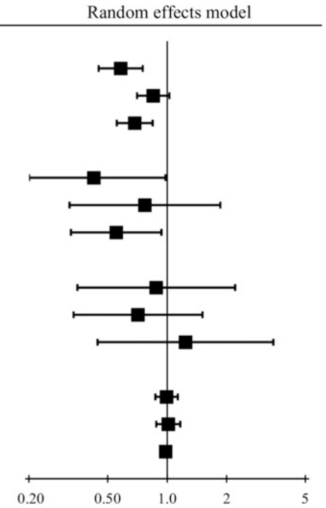

FIG. 4. Pooled hazard ratios for all-cause mortality, cardiovascular mortality, MACE, and cancer mortality outcomes. ATD, antithyroid drugs; RAI, radioactive iodine; MACE, major adverse cardiovascular event; HR, hazard ratio; CI, confidence interval.

levothyroxine replacement without delay. This prolonged unpredictable thyroid status after RAI may contribute to worse outcomes than after thyroid surgery. However, no significant differences were observed between RAI and ATD-treated patients.

Whether RAI increases cancer risk and cancer mortality has been debated. We found no significant difference in cancer mortality across the three treatments. Similarly, a recent meta-analysis of 12 studies involving 479,452 patients reported no increased cancer risk in RAI-treated versus nonexposed patients with hyperthyroidism, although a linear dose-response relationship with solid cancer mortality was observed.[9]

Given consistent evidence showing lower all-cause and cardiovascular mortality with surgery compared to ATD, it may be time to re-evaluate the role of long-term ATD therapy, although more studies are needed. ATD remains the preferred treatment for hyperthyroidism worldwide, largely driven by clinicians' and patients' hopes for Graves' disease remission without permanent hypothyroidism.[8] Concerns about the need for lifelong thyroid hormone replacement, radiation exposure, surgical complications, and visible neck scarring have further reinforced patients' preference.[11] Despite a relatively low remission rate of 30%–50% after 12–18 months of ATD, prolonged use beyond 5 years may raise remission rates to 80%–85%.[1] The safety profile of long-term ATD use is supported by emerging evidence,[15] and its ease of administration makes it an attractive option. As such, ATD continues to be the primary treatment for persistent or relapsed hyperthyroidism.[8] However, patients and clinicians should be aware that achieving remission may require years of therapy, and the hope of avoiding hypothyroidism may not materialize as long-term real-world data suggest that only 30%–40% of Graves' disease patients will ultimately remain euthyroid without the need for levothyroxine.[8]

Table 2. P-Score of Studied Outcomes

Outcome	*Treatment*	P-*score*	*Rank*
All-cause mortality	Surgery	0.9999	1
	RAI	0.4778	2
	ATD	0.0223	3
Cardiovascular mortality	Surgery	0.9821	1
	RAI	0.3661	2
	ATD	0.1517	3
MACE	RAI	0.7353	1
	Surgery	0.4751	2
	ATD	0.2896	3
Cancer mortality	Surgery	0.6421	1
	ATD	0.5116	2
	RAI	0.3464	3

ATD, antithyroid drugs; RAI, radioactive iodine; MACE, major adverse cardiovascular events.

Although this is not the first network meta-analysis evaluating outcomes in hyperthyroidism treated with ATD, RAI, or surgery,[34,35] our study is the first to specifically address cardiovascular mortality, MACE, and cancer mortality. We conducted a more thorough and comprehensive literature search and full-text review than prior studies, carefully assessing study populations and designs. Several studies from prior reviews were not included due to misalignment with the outcomes of interest in this study. A summary of excluded studies is provided in Supplementary Table S4. Issa et al. included only three studies assessing arrhythmia, atrial fibrillation, heart failure, and hypertension.[34] Although Liu et al. included 22 studies, 12 focused on Graves' ophthalmopathy.[35] Their mortality outcomes combined heterogeneous endpoints, including all-cause, cardiovascular, and cancer mortality, and they double-counted Ron et al. (1998) and Kitahara et al. (2020), which analyzed the same CTTFUS cohort.[20,21] In contrast, our analysis included four additional relevant studies not captured by Liu et al., improving our review.[17,22,23,27] Among these four, the cohort published by our group is by far the largest, including 114,062 patients, and was designed to minimize immortal time bias.[17] Therefore, our systematic review and network meta-analysis offers the most comprehensive evidence on the long-term outcomes of hyperthyroidism treatments.

Limitations

This study has several limitations, primarily due to the retrospective observational nature of included studies. Absence of RCTs may limit the strength of causal inferences and precision of pooled effect size. Selection bias related to baseline health status may have influenced comparative outcomes across treatment groups. We observed substantial heterogeneity in studies assessing all-cause mortality, cardiovascular mortality, and MACE, likely due to differences in methodological quality, patient populations, and adjustment strategies. As a result, the estimated relative risks may differ from absolute risks observed in RCTs. Variation in adjusted confounders, such as lack of adjustment for baseline FT4 and TSH levels,[17,24,25] duration and degree of control of hyperthyroidism,[17,24,25] comorbidities,[21] and concomitant medications[28,29] in some studies, and absence of individual-level data limited our ability to address residual confounding. The limited number of studies evaluating surgical type, surgical timing, and post-RAI thyroid status precluded subgroup analyses of these important clinical questions. Finally, most studies included mixed hyperthyroidism etiologies without stratification by underlying cause. Only Boaelaert et al. and Okosieme et al. specifically included patients with Graves' disease confirmed by thyrotropin receptor antibody positivity.[7,23] While Liu et al. described their cohort as having Graves' disease, the ICD-9 codes used (242.00/242.01) indicate toxic diffuse goiter and they did not confirm Graves' disease with serology.[16,27] Therefore, our findings are most applicable to patients with hyperthyroidism due to Graves' disease or toxic nodules but should be interpreted within the context of individual clinical scenarios.

Conclusions

In this network meta-analysis, surgery for hyperthyroidism was associated with significantly decreased risks of all-cause mortality and cardiovascular mortality compared with ATD and RAI. No significant differences were found among the three treatment modalities for MACE and cancer mortality. Findings should be interpreted with caution given the inherent limitations of retrospective observational studies, including heterogeneous patient populations, potential selection bias, and the limited number of eligible studies. Further large-scale observational studies and RCTs are needed to strengthen the evidence base for guiding hyperthyroidism treatment decisions. Large-scale prospective studies and individual participant data meta-analyses are needed to strengthen the evidence base for guiding treatment decisions in hyperthyroidism.

Acknowledgments

The authors thank Kashif M. Munir, MD, from the Division of Endocrinology, Diabetes and Nutrition, University of Maryland School of Medicine, Baltimore, Maryland, USA, for providing access to Covidence, the systematic review software.

Authors' Contributions

C.C.P.: Conceptualization (lead), methodology (equal), writing—original draft preparation (lead), visualization (lead), project administration (lead). B.R.S.: Validation (equal), data curation (support), writing—original draft preparation (support), visualization (support). D.Y.: Resources (lead), software (support), data curation (equal). H.K.H.: Methodology (support), formal analysis (support), writing—review and editing (support). C.C.L.: Conceptualization (support), project administration (support), resource (support), writing—review and editing (support). P.P.L.: Methodology (Equal), software (lead), validation (equal), formal analysis (lead), data curation (equal), project administration (support), writing—original draft preparation (support), visualization (support). E.N.P.: Conceptualization (support), methodology (support), writing—review and editing (lead), supervision (lead).

Author Disclosure Statement

C.C.P. reported receiving personal fees as a speaker for Mounjaro from Eli Lilly Taiwan and serving on the American Association of Clinical Endocrinology Management of Dyslipidemia Guideline Update Task Force outside the submitted work. E.N.P. reported receiving personal fees from the National Dairy Council and Merk China outside the submitted work and holding membership in the Management Council of the Iodine Global Network and the American Thyroid Association's thyroid in pregnancy clinical guidelines task force.

Funding Information

No funding was received for this study.

Supplementary Material

Supplementary Data

References

1. Lee SY, Pearce EN. Hyperthyroidism: A review. JAMA 2023;330(15):1472–1483; doi: 10.1001/jama.2023.19052
2. Chaker L, Cooper DS, Walsh JP, et al. Hyperthyroidism. Lancet 2024;403(10428):768–780; doi: 10.1016/s0140-6736(23)02016-0
3. Jabbar A, Pingitore A, Pearce SH, et al. Thyroid hormones and cardiovascular disease. Nat Rev Cardiol 2017;14(1):39–55; doi: 10.1038/nrcardio.2016.174
4. Sohn SY, Lee E, Lee MK, et al. The Association of Overt and Subclinical Hyperthyroidism with the risk of cardiovascular events and cardiovascular mortality: Meta-analysis and systematic review of cohort studies. Endocrinol Metab (Seoul) 2020;35(4):786–800; doi: 10.3803/EnM.2020.728
5. Dekkers OM, Horváth-Puhó E, Cannegieter SC, et al. Acute cardiovascular events and all-cause mortality in patients with hyperthyroidism: A population-based cohort study. Eur J Endocrinol 2017;176(1):1–9; doi: 10.1530/eje-16-0576
6. Lillevang-Johansen M, Abrahamsen B, Jørgensen HL, et al. Excess mortality in treated and untreated hyperthyroidism is related to cumulative periods of low serum TSH. J Clin Endocrinol Metab 2017;102(7):2301–2309; doi: 10.1210/jc.2017-00166
7. Okosieme OE, Taylor PN, Evans C, et al. Primary therapy of Graves' disease and cardiovascular morbidity and mortality: A linked-record cohort study. Lancet Diabetes Endocrinol 2019;7(4):278–287; doi: 10.1016/S2213-8587(19)30059-2
8. Villagelin D, Cooper DS, Burch HB. A 2023 international survey of clinical practice patterns in the management of

graves disease: A decade of change. J Clin Endocrinol Metab 2024;109(11):2956–2966; doi: 10.1210/clinem/dgae222
9. Shim SR, Kitahara CM, Cha ES, et al. Cancer risk after radioactive iodine treatment for hyperthyroidism: A systematic review and meta-analysis. JAMA Netw Open 2021; 4(9):e2125072; doi: 10.1001/jamanetworkopen.2021.25072
10. Kim BW. Does radioactive iodine therapy for hyperthyroidism cause cancer? J Clin Endocrinol Metab 2022;107(2): e448–e57; doi: 10.1210/clinem/dgab700
11. Daniels GH, Ross DS. Radioactive iodine: A living history. Thyroid 2023;33(6):666–673; doi: 10.1089/thy.2022.0344
12. van Kinschot CMJ, Soekhai VR, de Bekker-Grob EW, et al. Preferences of patients and clinicians for treatment of Graves' disease: A discrete choice experiment. Eur J Endocrinol 2021;184(6):803–812; doi: 10.1530/EJE-20-1490
13. Ross DS, Burch HB, Cooper DS, et al. 2016 American Thyroid Association guidelines for diagnosis and management of hyperthyroidism and other causes of thyrotoxicosis. Thyroid 2016;26(10):1343–1421; doi: 10.1089/thy.2016.0229
14. Azizi F, Amouzegar A, Tohidi M, et al. Increased remission rates after long-term methimazole therapy in patients with graves' disease: Results of a randomized clinical trial. Thyroid 2019;29(9):1192–1200; doi: 10.1089/thy.2019.0180
15. Azizi F, Malboosbaf R. Safety of long-term antithyroid drug treatment? A systematic review. J Endocrinol Invest 2019;42(11):1273–1283; doi: 10.1007/s40618-019-01054-1
16. Liu X, Wong CKH, Chan WWL, et al. Outcomes of graves' disease patients following antithyroid drugs, radioactive iodine, or thyroidectomy as the first-line treatment. Ann Surg 2021; 273(6):1197–1206; doi: 10.1097/SLA.0000000000004828
17. Peng CC, Lin YJ, Lee SY, et al. MACE and hyperthyroidism treated with medication, radioactive iodine, or thyroidectomy. JAMA Netw Open 2024;7(3):e240904; doi: 10.1001/jamanetworkopen.2024.0904
18. Hutton B, Salanti G, Caldwell DM, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: Checklist and explanations. Ann Intern Med 2015; 162(11):777–784; doi: 10.7326/M14-2385
19. Greenland S, Longnecker MP. Methods for trend estimation from summarized dose-response data, with applications to meta-analysis. Am J Epidemiol 1992;135(11):1301–1309; doi: 10.1093/oxfordjournals.aje.a116237
20. Ron E, Doody MM, Becker DV, et al. Cancer mortality following treatment for adult hyperthyroidism. Cooperative Thyrotoxicosis Therapy Follow-up Study Group. JAMA 1998;280(4):347–355; doi: 10.1001/jama.280.4.347
21. Kitahara CM, Preston DL, Sosa JA, et al. Association of Radioactive Iodine, Antithyroid Drug, and Surgical Treatments with solid cancer mortality in patients with hyperthyroidism. JAMA Netw Open 2020;3(7):e209660; doi: 10.1001/jamanetworkopen.2020.9660
22. Goldman MB, Maloof F, Monson RR, et al. Radioactive iodine therapy and breast cancer. A follow-up study of hyperthyroid women. Am J Epidemiol 1988;127(5):969–980; doi: 10.1093/oxfordjournals.aje.a114900
23. Boelaert K, Maisonneuve P, Torlinska B, et al. Comparison of mortality in hyperthyroidism during periods of treatment with thionamides and after radioiodine. J Clin Endocrinol Metab 2013;98(5):1869–1882; doi: 10.1210/jc.2012-3459
24. Giesecke P, Frykman V, Wallin G, et al. All-cause and cardiovascular mortality risk after surgery versus radioiodine treatment for hyperthyroidism. Br J Surg 2018;105(3): 279–286; doi: 10.1002/bjs.10665
25. Gronich N, Lavi I, Rennert G, et al. Cancer risk after radioactive iodine treatment for hyperthyroidism: A cohort study. Thyroid 2020;30(2):243–250; doi: 10.1089/thy.2019.0205
26. Hoffman DA, McConahey WM, Diamond EL, et al. Mortality in women treated for hyperthyroidism. Am J Epidemiol 1982; 115(2):243–254; doi: 10.1093/oxfordjournals.aje.a113296
27. Liu X, Wong CKH, Chan WWL, et al. Long-term outcome of patients treated with antithyroid drugs, radioactive iodine or surgery for persistent or relapsed Graves' disease. Br J Surg 2022;109(4):381–389; doi: 10.1093/bjs/znab474
28. Ryödi E, Metso S, Jaatinen P, et al. Cancer incidence and mortality in patients treated either with RAI or thyroidectomy for hyperthyroidism. J Clin Endocrinol Metab 2015; 100(10):3710–3717; doi: 10.1210/jc.2015-1874
29. Ryödi E, Metso S, Huhtala H, et al. Cardiovascular morbidity and mortality after treatment of hyperthyroidism with either radioactive iodine or thyroidectomy. Thyroid 2018; 28(9):1111–1120; doi: 10.1089/thy.2017.0461
30. Cohen O, Ronen O, Khafif A, et al. Revisiting the role of surgery in the treatment of Graves' disease. Clin Endocrinol (Oxf) 2022;96(6):747–757; doi: 10.1111/cen.14653
31. Bartalena L, Burch HB, Burman KD, et al. A 2013 European survey of clinical practice patterns in the management of Graves' disease. Clin Endocrinol (Oxf) 2016;84(1):115–120; doi: 10.1111/cen.12688
32. Parameswaran R, de Jong MC, Kit JLW, et al.; Asian Graves Consortium Study. 2021 Asia-Pacific Graves' Disease Consortium Survey of clinical practice patterns in the management of Graves' disease. Endocrine 2023;79(1): 135–142; doi: 10.1007/s12020-022-03193-7
33. Franklyn JA, Sheppard MC, Maisonneuve P. Thyroid function and mortality in patients treated for hyperthyroidism. JAMA 2005;294(1):71–80; doi: 10.1001/jama.294.1.71
34. Issa PP, Hussein M, Omar M, et al. Cardiovascular health by Graves' disease management modality – surgery versus radioactive iodine versus antithyroid medications: A network meta-analysis. J Surg Res 2023;283:266–273; doi: 10.1016/j.jss.2022.10.043
35. Liu X, Wong CKH, Wu T, et al. A comparison of cardiovascular disease, cancer, mortality, and Graves' ophthalmopathy following treatment for hyperthyroidism: A Bayesian network meta-analysis. World J Surg 2024;48(2):393–407; doi: 10.1002/wjs.12066

Address correspondence to:
Elizabeth N. Pearce, MD, MSc
Section of Endocrinology, Diabetes, Nutrition and Weight Management
Boston University Chobanian & Avedisian School of Medicine
Boston, MA
USA

E-mail: Elizabeth.Pearce@bmc.org

Peter Pin-Sung Liu, PhD
Population Health Data Center
National Cheng Kung University
Tainan
Taiwan

E-mail: psliu520@gmail.com