
Peer reviewed version
License (if available):
CC BY-NC-ND
Link to published version (if available):
10.1016/S2213-8587(19)30059-2

Link to publication record in Explore Bristol Research
PDF-document

This is the accepted author manuscript (AAM). The final published version (version of record) is available online via The Lancet at https://doi.org/10.1016/S2213-8587(19)30059-2. Please refer to any applicable terms of use of the publisher.

**University of Bristol - Explore Bristol Research**

**General rights**

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
http://www.bristol.ac.uk/pure/about/ebr-terms
Primary therapy of Graves’ disease and cardiovascular morbidity and mortality: a linked-record cohort study

Onyebuchi E Okosieme¹,², Peter N Taylor¹, Carol Evans³, Dan Thayer⁴, Aaron Chai¹, Ishrat Khan¹, Mohd S Draman¹, Brian Tennant⁵, John Geen⁶,⁷, Adrian Sayers⁷, Robert French¹, John H Lazarus¹, Lakdasa D Premawardhana¹,⁸, and Colin M Dayan¹

1. Thyroid Research Group, School of Medicine, Cardiff University
2. Diabetes Department, Prince Charles Hospital, Cwm Taf University Health Board
3. Department of Medical Biochemistry & Immunology, University Hospital of Wales, Cardiff
4. SAIL Databank, School of Medicine, Swansea University
5. Clinical Biochemistry Department, Cwm Taf University Health Board
6. Faculty of Life Sciences and Education, University of South Wales
7. Department of Social and Community Medicine, University of Bristol, Bristol BS8 2BN, United Kingdom.
8. Section of Endocrinology, Department of Medicine, YYF Hospital, Aneurin Bevan University Health Board

Key words: Graves’ disease, hyperthyroidism, radioiodine therapy, thyroidectomy, mortality, major adverse cardiovascular events

Word count: 3542
Tables: 2
Figures: 5
References: 33

Correspondence: Dr Onyebuchi E Okosieme, Diabetes Department, Prince Charles Hospital, Cwm Taf University Health Board, Gurnos Estate, Merthyr Tydfil, CF47 9DT, Telephone 01685728353, Fax 01685728448
ABSTRACT

BACKGROUND: Graves’ disease is routinely treated with antithyroid drugs (ATD), radioactive-iodine (RAI), or surgery but it is uncertain whether the choice of initial therapy influences long-term outcomes. We evaluated cardiovascular morbidity and mortality according to the modality and effectiveness of primary therapy in Graves’ disease.

METHODS: The study was conducted through linked datasets within the All-Wales Secure Anonymised Information Linkage (SAIL) Databank. Graves’ disease patients were identified from a regional TSH-receptor-antibody test register, (n=4,189, female 82%, 1998–2013) and matched by age and sex to a control population in SAIL (n=16,756). Patients were grouped by treatment within one-year of diagnosis as: (1) ATD (n=3587), (2) RAI with resolved hyperthyroidism (RAI-Group-A, n=250), and (3) RAI with unresolved hyperthyroidism (RAI-Group-B, n=182). One-year landmark Kaplan-Meier and Cox regression models were used to analyse the association of treatment with all-cause mortality and major adverse cardiovascular events (MACE, myocardial infarction, heart failure, ischaemic stroke, or death). The relationship between FT4 concentration and outcomes was analysed using restricted cubic-spline regression models.

FINDINGS: Overall, patients had increased mortality compared to controls (HR 1·23, 95%CI 1·06, 1·42). Compared to ATD-treated patients, mortality was reduced in RAI-Group-A (HR 0·50, 95%CI 0·29, 0·85) but not RAI-Group-B (HR 1·51, 95%CI 0·96, 2·37). Persistently low-TSH at 1-year was associated with increased mortality independent of treatment modality (HR 1·55, 95%CI 1·08-2·24). Spline-regressions demonstrated a positive non-linear relationship between 1-year-FT4 and outcomes.
**INTERPRETATION:** Regardless of therapy modality, early and effective control of Graves’ hyperthyroidism is associated with improved survival compared to less effective control.

**FUNDING:** National Institute for Social Care and Health Research (NISCHR), Wales
RESEARCH IN CONTEXT

Evidence before this study

We searched PubMed for articles published up to October 2018, without language restrictions, that addressed mortality outcomes in relation to treatment modality in patients with hyperthyroidism. Our search terms were (“hyperthyroidism” OR “thyrotoxicosis” OR “Graves’ disease”) AND (“mortality” OR “survival”) AND (“treatment” OR “antithyroid drugs” OR “thionamides” OR “radioactive iodine” OR “thyroidectomy”). Although several studies reported excess mortality in radioiodine-treated cohorts, these studies did not include patients treated with antithyroid drugs which currently represents the predominant therapy modality for Graves’ disease worldwide. Only one study to date has compared hyperthyroidism mortality according to treatment with antithyroid drugs or radioiodine. This study showed survival advantages following radioiodine-induced hypothyroidism compared to during antithyroid drug treatment. However, the study comprised a heterogeneous hyperthyroid cohort that was underpowered to show treatment effects in the Graves’ disease subset of 345 patients.

Added value of this study

To the best of our knowledge ours is the first study to show that early and complete resolution of hyperthyroidism in patients with Graves’ disease is associated with improved survival compared to less effective control. We show that survival in patients with Graves’ disease is associated with the rapid restoration of normal thyroid status rather than therapy modality per se. Using flexible spline regression modelling we demonstrate a positive association between FT4 levels after initial therapy and subsequent cardiovascular morbidity and mortality.
Implications of all the available evidence

Our findings emphasise the importance of prompt elimination of hyperthyroidism and maintenance of normal thyroid status in the care of Graves’ disease patients. The choice of therapy should be driven by the prospects of successful resolution of hyperthyroidism and therefore, early definitive therapy should be offered to patients unlikely to achieve remission with antithyroid drugs alone.
INTRODUCTION

Graves’ disease is the most common cause of hyperthyroidism in iodine-sufficient countries, affecting 0.5–1.0% of the population\textsuperscript{1,2}. Hyperthyroidism in Graves’ disease results from unregulated stimulation of the thyrotropin (TSH) receptor by autoreactive TSH-receptor antibodies (TRAbs) which are detectable in virtually every patient with the disease\textsuperscript{3}. Patients with uncorrected hyperthyroidism have an increased mortality risk and suffer significant cardiovascular morbidity including cardiac arrhythmias, strokes and heart failure\textsuperscript{4,5}. Three treatment modalities for Graves’ disease, namely thionamide drugs, radioactive iodine, and thyroidectomy have been available for decades, all with established safety and efficacy records. However, the approach to therapy is highly variable and largely dictated by regional preferences\textsuperscript{2,6}. Primary thionamide therapy is favoured by the European Thyroid Association\textsuperscript{7}, while the American Thyroid Association endorses any of the modalities depending on individual patient characteristics\textsuperscript{8}. In practice radioiodine is traditionally preferred in North America\textsuperscript{9} although antithyroid drugs are increasingly used and are now the most common treatment for Graves’ disease in the United States \textsuperscript{10}. In contrast primary radioiodine therapy is less popular in Europe and Asia and remains restricted to patients who relapse following 12-18 months of thionamides\textsuperscript{11-12}.

These variations in treatment are in part driven by lack of evidence regarding mortality risk with respect to Graves’ disease management. Randomised controlled trials on mortality outcomes in Graves’ disease are lacking and are unlikely to be undertaken given the logistic implications. Although observational studies have addressed mortality end-points in patients with hyperthyroidism\textsuperscript{4,12-18}, only a few studies specifically distinguished Graves’ disease from other causes of hyperthyroidism\textsuperscript{17,19-23}. This distinction is relevant given that Graves’ disease patients are younger and
express a different cardiovascular phenotype from those with thyroid nodules\textsuperscript{2,5,20,22}. In addition, unlike patients with toxic nodules who require definitive treatment for cure, Graves’ disease patients have a choice of all three treatments, underpinning the need for treatment-specific outcome data.

Although previous cohort studies have reported excess mortality in radioiodine-treated patients\textsuperscript{15,17,18,23,24}, these cohorts excluded patients treated with thionamides which currently represents the predominant therapy modality for Graves’ disease worldwide\textsuperscript{2}. Only one study to date has compared mortality according to thionamide versus radioiodine treatment\textsuperscript{19}. The study by Boelaert et al observed reduced mortality after radioiodine induced hypothyroidism compared to during thionamide therapy. However, the study comprised a heterogeneous hyperthyroidism clinic cohort (n=1036) that was underpowered to show treatment effects in the Graves’ disease subset (n=345)\textsuperscript{19}. Thus, it remains unclear whether primary radioiodine therapy offers survival advantages over thionamides and hence whether current Graves’ disease treatment strategies are optimal. Our objective in the present study was to evaluate cardiovascular morbidity and mortality according to the modality and effectiveness of primary therapy in a large cohort of patients with Graves’ disease.

**METHODS**

*Data Sources and study population*

The study was conducted through linked datasets within the Secure Anonymised Information Linkage (SAIL) Databank. SAIL is a nationwide repository of routinely collected health and social care data in Wales run by the Health Informatics Research Unit of Swansea University. The databank contains over two billion anonymised person-based records linked to hospital admission and primary care data\textsuperscript{25}. All National Health Service (NHS) users in Wales are identified by a unique 10-digit
identifier and anonymity of data within the databank is ensured by multi-party encryption through a split-file data management system\textsuperscript{26}. For this study, we created a dataset of patients with hyperthyroidism, diagnosed between January 1998 and December 2013 in clinics across three health boards serving an estimated 1·4 million residents in South Wales. We included all patients who could be linked to the Welsh Demographic Service, a population register of individuals registered with a general practitioner (GP) or receiving healthcare in Wales. Additional outcomes were obtained from primary care data held within SAIL, covering about 75\% of the population, and the Patient Episode Database for Wales (PEDW), a dataset of all inpatient admissions in Wales.

Patients with Graves’ disease were identified from electronic records of TRAb tests performed at the biochemistry department of the University Hospital of Wales, Cardiff, which served as the regional TRAb-assay laboratory during the period covered by the study. TRAbs are sensitive and specific markers of Graves’ disease\textsuperscript{27} and are routinely included in the diagnostic work-up of Graves’ disease in our catchment clinics. Patients were defined as Graves’ Hyperthyroidism if they had a positive TRAb test together with evidence of hyperthyroidism from either: (a) low-TSH in laboratory records, (b) record of hyperthyroidism in GP or hospital datasets (ICD-10 E05, read codes C02), or (c) record of thionamide prescription in GP data (carbimazole, read codes fa1 and propylthiouracil, fa2). The patient inclusion flow chart is presented in figure 1. For a background control population, we identified a sample of individuals in the Welsh Demographic Service, matched to cases on birth week and sex, at a ratio of 4 control subjects per patient (n=16,756). Individuals with hyperthyroidism on our loaded datasets were excluded from the controls.
**Treatment and outcome data**

Iodine-131 treatment data was extracted from radioiodine treatment logs in nuclear medicine departments serving the study population. Each department holds statutory records of all radioisotope treatment including isotope type, dose, and date of administration. Information on thyroidectomy was obtained from PEDW using the Office of Population censuses and Surveys version 4 (OPCS-4) classification code for thyroidectomy, B081-B091. Outcome events were recorded from PEDW using the relevant ICD-10 codes for atrial fibrillation (I48), acute myocardial infarction (I21-I23), acute coronary syndrome (I20), heart failure (I50), ischaemic stroke (I63, I65-I68), haemorrhagic stroke (I60-I62), diabetes mellitus (E10-E14), thyroid cancer (C73), and breast cancer (C50). Mortality information was obtained from the Office for National Statistics (ONS).

**Laboratory assays**

TRAbs were performed with second generation radioimmunoassays while thyroid hormones (FT4 and TSH) were performed using several different immunoassays. The laboratory assays, reference ranges and performance measures are detailed in the supplementary appendix and supplementary table 1.

**Statistical analysis**

Baseline data is summarised as means (standard deviation, SD) for normally distributed values or median (interquartile range, IQR) for non-normal data. Differences between groups were analysed with chi-squared test, 2-tailed t-test or Mann-Whitney tests as appropriate. The primary study outcome was all-cause mortality while secondary outcome was major adverse cardiovascular events (MACE), a composite of acute myocardial infarction, ischaemic stroke, heart failure, or death. Patients were grouped according to primary treatment using a landmark approach to
avoid immortal time bias. The landmark was set at one-year post-diagnosis which was considered sufficient time for primary therapy without incurring excessive data losses from pre-landmark censoring.

Treatment groups comprised: (1) ATD, patients who received only antithyroid drugs in the first year of diagnosis, (2) RAI-Grp-A, patients who received radioiodine and achieved hyperthyroidism control by one-year, and (3) RAI-Grp-B, patients who received radioiodine but had not achieved hyperthyroidism control by one-year. We excluded patients if they were censored before one-year (n=93) or had thyroidectomy in the first-year due to small numbers (78 patients, 2 deaths). Control of hyperthyroidism after radioiodine was defined as levothyroxine initiation, or TSH above the reference limit, or persistent euthyroidism (>6 months), each without antithyroid drug use. The diagnosis date was the date of the first positive TRAb or low-TSH, whichever was recorded earlier, and follow-up was from diagnosis until event or study end.

We analysed outcomes according to primary treatment using Kaplan-Meier curves and Cox regression models. First, we compared event rates in patients and population controls using the logrank test and an adjusted Cox regression model to account for baseline comorbidity. The analysis was undertaken in the whole cohort and stratified by primary treatment group, with the corresponding matched population for each group as controls. In a second approach, we used a Cox regression model to compare event risk between treatment groups with the antithyroid drug group as reference category. Models were adjusted for 5-year age bands, gender, diagnosis-year, baseline TRAb concentration, and comorbidity scores, using a modification of the Charlson Comorbidity Index (supplementary table 2). Stratified analysis was undertaken by age (>50 vs. ≤50 years) and co-morbidity score (0 vs ≥1) with interaction terms for sub-
groups. In sensitivity analysis, we sequentially excluded pregnant patients (n=63) and patients with borderline TRAb levels (n=266) at diagnosis. The validity of the proportional hazard assumption was examined by Schoenfeld residuals test and visual inspection of residuals plots against time.

As an alternative means of handling time bias, treatment and thyroid status were incorporated as time-dependent covariates in a multivariable Cox proportional hazard regression model in which patients contributed person-time as follows: (1) ATD, from diagnosis to definitive treatment with radioiodine or thyroidectomy; (2) RAI without hyperthyroidism control, from radioiodine therapy until control of hyperthyroidism, (3) RAI with hyperthyroidism control, following radioiodine control of hyperthyroidism; (4) Thyroidectomy, following thyroidectomy.

To directly evaluate the impact of thyroid status, we analysed outcomes according to thyroid function after one-year of treatment. We derived a one-year FT4 and TSH from time-weighted averages of all available tests at 9–12 months excluding patients censored before one-year. TSH was entered in a Cox regression model as low, normal, or high-TSH with the normal-TSH as reference category. Models were adjusted for age, gender, comorbidity, baseline TRAb, FT4 and TSH, and then stratified by primary treatment modality based on treatment received by the test date. For FT4, we used restricted cubic splines to model a possible non-linear relationship between one-year FT4 and outcomes. Four equally-spaced knots were set at percentiles 5, 35, 65 and 95 according to Harrell’s recommendation. To account for differences in assay methods, FT4 was converted to multiples of the upper limit of the assay reference range (ULN). Restricted cubic splines were also used to model the longitudinal FT4 change in in the first 24-months of treatment with cubic knots set at 0, 3, 6, 9, 12, and 18 months.
Serial thyroid function tests were missing in 29% of patients. The characteristics of patients with thyroid tests were comparable to those with missing tests except that patients with tests were younger (mean age 47 vs 49 years, P<0.001), and missing tests were more frequent in patients treated with antithyroid drugs compared to radioiodine or surgery (37% vs 3% vs 16%, P<0.001) (supplementary appendix, supplementary table 3). Thus, we addressed missing tests in the one-year TSH analysis using multiple variable imputation by chained equations\textsuperscript{29} under the missing at random assumption (supplementary appendix). We generated 50 imputed datasets with 500 iterations and fitted Cox proportional models within each dataset after which estimates were pooled according to Rubin’s rules. Sensitivity analysis using a complete case analysis gave similar results, so we report results from the imputed datasets while the complete case analysis is presented in the supplementary appendix. Statistical analysis was performed via remote desktop with the SAIL Databank using Stata version 16 for Windows (Stata Corp., College Station, TXS, USA).

**Regulatory Approvals**

The study was registered with the Research and Development Office of Cwm Taf Health Board and approval for data linkage with individual health boards was obtained (reference CT/688/140049/16). Approval for access to anonymised data in the SAIL databank was granted by the SAIL Information Governance Review panel (project reference, 0233). Ethical permission is granted for conducting data analysis in the SAIL Databank and separate ethics committee review was not required for this study.
Role of the funding source

The funding source had no role in study design, data analysis, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

RESULTS

The final study cohort comprised 4,189 TRAb-positive patients with Graves’ disease and 16,756 controls, with follow-up ranging from 6 months to 16·8 years in the SAIL databank. Baseline characteristics are presented in table 1. 82% of patients were females with peak age 30–60 years. Overall, antithyroid drug treatment alone was used in 74% of patients while 20% and 6% received radioiodine or thyroidectomy respectively. Of the whole cohort, 4,019 patients with at least one-year follow-up were eligible for inclusion in the landmark analysis including 3,587 patients treated with antithyroid drugs (ATD) alone in the first-year, 250 radioiodine-treated patients with resolved hyperthyroidism by one-year (RAI-Grp-A), and 182 radioiodine-treated patients with unresolved hyperthyroidism by one-year (RAI-Grp-B). The characteristics of the landmark cohort is presented in Supplementary table 4. As in the full cohort, radioiodine treated patients were older than medically-treated patients (mean age 53 vs 47 years) and had more severe disease at diagnosis in terms of higher baseline levels of FT4 and TRAbs than ATD patients. Although both radioiodine groups were well-balanced in baseline characteristics and iodine-131 dose (median 555 Mbq), RAI-Grp-A patients received earlier iodine-131 treatment than RAI-Grp-B (median time to treatment of 108 vs 188 days).

Overall, 228 deaths (5·4%) were recorded in patients compared to 765 (4·6%) in controls. Crude and adjusted Cox regression models for mortality and other events are presented in table 2. Patients had increased mortality (HR 1·23, 95%CI 1·06, 1·42)
and MACE (HR 2.47, 95%CI 2.16, 2.81) compared to controls and also had increased risks for atrial fibrillation, congestive cardiac failure, and ischaemic stroke (table 2). Figure 2 shows survival curves in patients versus controls, presented for the whole cohort and stratified by primary treatment groups. The adjusted Cox regression model showed an increase in mortality and MACE in all treatment groups apart from RAI-Grp-A patients (HR 0.65, 95%CI 0.38, 1.13). Logrank tests gave consistent results with the adjusted model except for mortality in RAI-Grp-B (figure 2d) and MACE in RAI-Grp-A (figure 2g) which failed to reach significance by logrank but were significant in the adjusted model. The results remained unchanged when all treatment groups and the control population were included in an adjusted Cox regression model with a shared frailty term for case-control clusters.

Between-group comparisons are shown in figure 3. RAI-Grp-A patients had reduced risk of mortality (HR 0.50, 95% CI 0.29, 0.85) and MACE (HR 0.59, 95% CI 0.38, 0.92) compared to the ATD group. In contrast, RAI-Grp-B patients had an increase in MACE risk (HR 1.52, 95% CI 1.01, 2.28) with no survival advantage over ATD patients. Similar results were seen in sub-groups without comorbidity but not those with comorbidity (P for interaction 0.039). There was no interaction with age (>50 vs ≤50 years) and results remained unchanged after excluding pregnant patients or patients with borderline TRAbs (supplementary tables 5 and 6).

The results of the time-dependent multivariable analysis are presented in supplementary tables 7 and 8. The models incorporated serial thyroid hormones and treatment modality as time-dependent variables. Mortality and MACE risk increased with age, male gender, and co-morbidity but reduced with increasing year of diagnosis possibly reflecting temporal improvements in overall cardiovascular survival. Increased risk of mortality and MACE was observed with increasing levels of serial
FT4 but not serial TSH. With respect to treatment, reduced mortality and MACE risk was seen after radioiodine induced control of hyperthyroidism compared to during antithyroid drug therapy. In contrast reduced mortality was not seen after radioiodine therapy that had not controlled hyperthyroidism or after thyroidectomy although there was low power due to fewer thyroidectomies (supplementary table 7–8).

To evaluate the influence of thyroid status on outcomes we analysed the relationship of TSH at one-year with mortality and MACE using a landmark Cox regression model (figure 4). Of the patients with a one-year TSH, 898 (39%) had low-TSH while 429 (19%) had high-TSH. The fully adjusted model showed increased mortality and MACE risk with Low-TSH but not high-TSH (figure 4). This association was observed in the whole cohort as well as in the antithyroid drug group with no significant interaction with treatment (P value for interaction of 0.325 and 0.565 for mortality and MACE respectively, figure 4). Similar results were obtained from the imputation model (figure 3) and the complete case analysis (supplementary figure 1).

The relationship between FT4 and outcomes was modelled with restricted cubic splines which showed increasing hazard with FT4 concentrations above the reference range (figure 5). This relationship was seen in the whole cohort and in drug-treated patients but not radioiodine-treated patients (figure 5). Cubic splines were also used to model the longitudinal change in FT4 after diagnosis. All treatment groups showed a rapid drop in FT4 after diagnosis which was most pronounced and sustained in the successfully treated radioiodine group (supplementary figure 2).

**DISCUSSION**

We have used linked health records to evaluate all-cause mortality and cardiovascular event risk in patients with Graves’ disease and compared outcomes according to the modality and effectiveness of primary treatment. Compared to population controls,
Graves’ disease patients had a 23% increase in all-cause mortality and over twice the risk of a major cardiovascular event. Furthermore, early and successful treatment with radioiodine was associated with a 50% reduction in mortality compared to primary treatment with antithyroid drugs. Notably, survival advantages were not seen in radioiodine treated patients who failed to achieve control of hyperthyroidism. Patients with persistently low-TSH after one-year of treatment had a 55% increase in mortality even in patients treated with antithyroid drugs alone. Thus, for the first time, we show that mortality in patients with Graves’ disease is reduced by early and intensive correction of hyperthyroidism regardless of therapy modality.

Our findings confirm the increased mortality risk previously reported in some19,20,22 but not all17,30 hyperthyroidism outcome studies. However, only a few studies to date have specifically addressed mortality risk in relation to hyperthyroidism treatment. Previous large-scale cohort-studies suggested an excess of cardiovascular mortality in radioiodine-treated patients compared to the background population15,17,18,31. A 1980s United States thyroid clinic study observed no difference in mortality between radioiodine and thyroidectomy-treated women32 whereas more recent registry studies from Finland 24 and Sweden23 have reported survival advantages of thyroidectomy over radioiodine treatment. The interpretation of these data is clearly constrained by study designs that excluded thionamide-treated patients15,17,18,23,24,31,32, failed to account for thyroid status23,31,32, and did not control for bias from uneven time-intervals preceding definitive therapy15,17,18,31,32. In their Birmingham clinic study, Boelaert et al applied a time-dependent regression model that incorporated thyroid status to show survival benefits relative to thionamides following radioiodine-induced hypothyroidism19. However, their study was underpowered to show benefits in the Graves’ disease subset19.
In the present study, we have employed a study design that distinguished treatment effects from the effects of hyperthyroidism to show that the predictor of cardiovascular morbidity and mortality is the effectiveness of primary treatment rather than therapy modality *per se*. Our one-year TSH analysis shows that the excess mortality in Graves’ disease is driven by exposure to uncontrolled hyperthyroidism and that elimination of the hyperthyroid state whether with antithyroid drugs or radioiodine yields survival benefits. This is biologically plausible given the established association of hyperthyroidism with cardiovascular disease\(^4\) and is consistent with a recent Danish register-based study that reported excess mortality in association with cumulative periods of low-TSH in patients with treated and untreated hyperthyroidism\(^{14}\). Subsequent analyses of the Danish register also showed increased mortality with cumulative hypothyroidism exposure\(^{33}\) highlighting the need for strict surveillance of therapy to avoid periods of uncontrolled hypothyroidism or hyperthyroidism. As our data shows, persistent thyroid dysfunction is common during Graves’ disease therapy, with low or high TSH recorded in 58% of patients at one-year.

The choice of primary therapy in Graves’ disease is currently driven by regional preferences\(^{7,8}\). Our findings indicate that a greater value should now be placed on the rapid elimination of hyperthyroidism with maintenance of a normal-TSH as standard of care. Thus, patients with unfavourable characteristics for thionamide-induced relapse such as severe disease, or high TRAb levels should be offered primary radioiodine therapy. The use of radioiodine should be governed by treatment protocols designed to optimise treatment success through efficient patient preparation and administration of ablative doses of iodine-131. Primary definitive treatment will not be a priority for patients with mild disease and in such cases thionamides should be used at doses and duration that promote rapid and sustained control of hyperthyroidism.
The strength of our study is the unique dataset of 4,189 TRAb-positive Graves’ disease patients with long-term follow-up and robust and innovative data linkage. Ours is the first large-scale hyperthyroidism cohort study to exploit TRAbs as a specific Graves’ disease marker, thereby excluding patients with toxic nodules, thyroiditis, or non-thyroidal illness. By employing a landmark regression approach, we eliminated treatment time bias and additionally provided a clinically relevant time-point from which to evaluate primary therapy. Our study limitation is that the retrospective design could have introduced bias from undocumented confounders such as smoking, orbitopathy, physician treatment preference, patient compliance with thionamides, and unrecorded exposure to thyroid hormone excess. However, our adjustments for comorbidity and biochemical characteristics would to some extent have adjusted for these confounders. In keeping with conventional practice, patients with more severe baseline disease received radioiodine, but it is unlikely that this influenced our results given that both radioiodine-treated groups had similar baseline features yet strikingly divergent outcomes. Lastly, thyroid tests were missing in a proportion of patients. However, we minimised potential bias from this using principled imputation methods and robust sensitivity checks. Ultimately, potential confounders will only be fully addressed in well-designed, randomised controlled trials which are yet to be undertaken.

In conclusion, early and intensive control of hyperthyroidism in Graves’ disease is associated with improvements in long-term cardiovascular morbidity and mortality regardless of therapy modality. Greater emphasis should now be placed on ensuring prompt elimination of hyperthyroidism and maintenance of a normal thyroid status in the management of Graves’ disease.

Contributors
CMD and OO conceived and designed the study. OO performed the statistical analysis with contributions from CMD, RF, PNT, DT, and AS. OO wrote the first draft of the manuscript and all authors contributed to subsequent drafts and approved the manuscript before submission for publication. OO had full access to the study data and vouches for the integrity and accuracy of the data analysis.

Declaration of interests

OO reports grants from the National Institute for Social Care and Health Research, Wales, during the conduct of the study, and grants from Concordia International, outside the submitted work. DT reports grants from Bristol-Myers Squibb, grants from Pfizer, grants from Janssen-Cilag Ltd., and grants from Medical Ethics Pty., outside the submitted work. CMD reports advisory board membership of Apitope, advisory board membership of Sanofi-Genzyme, and lecture fees from Sanofi-Genzyme, outside the submitted work.

Acknowledgements

OO was supported for the study by an Academic Health Science Collaboration Fellowship of the National Institute of Social Care and Health Research Wales.

RESULTS

TABLE 1: Baseline characteristics of the whole cohort

TABLE 2: Hazard ratios for mortality, MACE, and other events

FIGURE 1: Study Flow chart

FIGURE 2: Kaplan Meier curves for mortality or MACE

FIGURE 3: Hazard ratios for mortality or MACE by primary treatment

FIGURE 4: Hazard ratios for mortality or MACE by one-year TSH

FIGURE 5: One-year FT4 and Mortality and MACE
SUPPLEMENTARY FIGURES AND TABLES

SUPPLEMENTARY FIGURE 1: Hazard ratios for mortality and MACE by one-year TSH: Complete Case Analysis

SUPPLEMENTARY FIGURE 2: FT4 profile post-diagnosis

SUPPLEMENTARY TABLE 1: Thyroid hormone assays

SUPPLEMENTARY TABLE 2: Modified Charlson Comorbidity Scores

SUPPLEMENTARY TABLE 3: Baseline characteristics in patients with missing and non-missing thyroid function tests

SUPPLEMENTARY TABLE 4: Characteristics of patients in the landmark cohort

SUPPLEMENTARY TABLE 5: Hazard ratios for Mortality and MACE by primary treatment: excluding pregnant patients

SUPPLEMENTARY TABLE 6: Hazard ratios for Mortality and MACE by primary treatment: excluding patients with borderline TRAb levels

SUPPLEMENTARY TABLE 7: Multivariable analysis for all-Cause mortality

SUPPLEMENTARY TABLE 8: Multivariable analysis for MACE
REFERENCES


<table>
<thead>
<tr>
<th></th>
<th>Population controls</th>
<th>All Patients</th>
<th>Treatment groups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n=16756</td>
<td>n=4189</td>
<td>Antithyroid drugs</td>
</tr>
<tr>
<td>Age, yrs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>48 ± 16</td>
<td>48 ± 16</td>
<td>48 ± 16</td>
</tr>
<tr>
<td>&lt;20</td>
<td>204 (1·2)</td>
<td>51 (1·2)</td>
<td>37 (1·2)</td>
</tr>
<tr>
<td>20-29</td>
<td>2176 (13·0)</td>
<td>544 (13·0)</td>
<td>402 (13)</td>
</tr>
<tr>
<td>30-39</td>
<td>3420 (20·4)</td>
<td>855 (20·4)</td>
<td>626 (20·2)</td>
</tr>
<tr>
<td>40-49</td>
<td>3688 (22·0)</td>
<td>922 (22·0)</td>
<td>667 (21·6)</td>
</tr>
<tr>
<td>50-59</td>
<td>3240 (19·3)</td>
<td>810 (19·3)</td>
<td>601 (19·4)</td>
</tr>
<tr>
<td>60-69</td>
<td>2340 (14·0)</td>
<td>585 (14·0)</td>
<td>443 (14·3)</td>
</tr>
<tr>
<td>&gt;70</td>
<td>1688 (10·1)</td>
<td>422 (10·1)</td>
<td>318 (10·3)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>13656 (81·5)</td>
<td>3414 (81·5)</td>
<td>2527 (81·7)</td>
</tr>
<tr>
<td>Male</td>
<td>3100 (18·5)</td>
<td>775 (18·5)</td>
<td>567 (18·3)</td>
</tr>
<tr>
<td>Comorbidity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>15453 (92·2)</td>
<td>3599 (85·9)</td>
<td>2647 (85·6)</td>
</tr>
<tr>
<td>Present</td>
<td>1300 (7·8)</td>
<td>590 (14·1)</td>
<td>447 (14·4)</td>
</tr>
<tr>
<td>TRAb IU/L</td>
<td>—</td>
<td>6·9 (3·1, 16·8)</td>
<td>6·0 (2·8, 13·6)</td>
</tr>
<tr>
<td>Thyroid status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FT4, pmol/L a</td>
<td>—</td>
<td>31·8 ± 22·0</td>
<td>29·5 ± 19·7</td>
</tr>
<tr>
<td>FT4 &gt;40 pmol/L a</td>
<td>—</td>
<td>719 (24)</td>
<td>390 (20)</td>
</tr>
<tr>
<td>Person years follow up</td>
<td>103374</td>
<td>25266</td>
<td>16481</td>
</tr>
</tbody>
</table>

**TABLE 1: Baseline characteristics of the whole cohort**

Data represents numbers (%), mean ± SD, or median (interquartile range).  
a, data available for 2962 patients, TRAb, TSH receptor antibody;  
c, P values are for trend across treatment categories. Patients and controls were matched for age.  
* Comorbidity was increased in patients compared to controls (P<0·001)
<table>
<thead>
<tr>
<th>Outcome</th>
<th>Number of cases (%)</th>
<th>Crude HR (95% CI)</th>
<th>P value</th>
<th>Adjusted HR (95 CI) b</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>All-cause Mortality</td>
<td>228 (5.4)</td>
<td>1.22 (1.05, 1.42)</td>
<td>0.008</td>
<td>1.23 (1.06, 1.42)</td>
<td>0.033</td>
</tr>
<tr>
<td>MACE c</td>
<td>329 (7.9)</td>
<td>2.18 (1.91, 2.48)</td>
<td>&lt;0.001</td>
<td>2.47 (2.16, 2.81)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Atrial Fibrillation</td>
<td>240 (5.7)</td>
<td>2.62 (2.23, 3.08)</td>
<td>&lt;0.001</td>
<td>2.67 (2.25, 3.15)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Acute Myocardial Infarction</td>
<td>38 (0.9)</td>
<td>1.26 (0.87, 1.81)</td>
<td>0.220</td>
<td>1.22 (0.84, 1.76)</td>
<td>0.287</td>
</tr>
<tr>
<td>Acute coronary syndrome</td>
<td>32 (0.8)</td>
<td>1.75 (1.15, 2.65)</td>
<td>0.008</td>
<td>1.65 (1.09, 2.51)</td>
<td>0.018</td>
</tr>
<tr>
<td>Congestive Cardiac Failure</td>
<td>109 (2.6)</td>
<td>1.85 (1.47, 2.31)</td>
<td>&lt;0.001</td>
<td>1.78 (1.42, 2.23)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Ischaemic Stroke</td>
<td>41 (1.0)</td>
<td>1.70 (1.17, 2.43)</td>
<td>0.005</td>
<td>1.60 (1.11, 2.30)</td>
<td>0.012</td>
</tr>
<tr>
<td>Haemorrhagic Stroke</td>
<td>14 (0.3)</td>
<td>1.53 (0.83, 2.83)</td>
<td>0.175</td>
<td>1.50 (0.81, 2.78)</td>
<td>0.196</td>
</tr>
<tr>
<td>Type 1 Diabetes</td>
<td>29 (0.7)</td>
<td>2.71 (1.68, 4.37)</td>
<td>&lt;0.001</td>
<td>2.53 (1.56, 4.08)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Type 2 Diabetes</td>
<td>170 (4.1)</td>
<td>1.46 (1.23, 1.74)</td>
<td>&lt;0.001</td>
<td>1.39 (1.16, 1.66)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>All Diabetes</td>
<td>199 (4.8)</td>
<td>1.57 (1.33, 1.85)</td>
<td>&lt;0.001</td>
<td>1.46 (1.24, 1.71)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Thyroid Cancer</td>
<td>8 (0.2)</td>
<td>6.52 (2.13, 19.93)</td>
<td>&lt;0.001</td>
<td>6.67 (2.18, 20.41)</td>
<td>0.001</td>
</tr>
<tr>
<td>Breast Cancer</td>
<td>38 (0.9)</td>
<td>1.08 (0.76, 1.55)</td>
<td>0.442</td>
<td>1.06 (0.74, 1.52)</td>
<td>0.798</td>
</tr>
</tbody>
</table>

**TABLE 2: Hazard ratios for mortality, MACE, and other events**

a Controls are matched for age and sex b Hazard ratios are derived from Cox proportional models adjusted for age, sex and comorbidity, c MACE, major adverse cardiovascular event comprising acute myocardial infarction, congestive cardiac failure, ischaemic cerebrovascular accident, or death
FIGURE 1: Study flow chart

TRAb, TSH receptor antibody; ATD, Antithyroid Drugs; RAI, radioactive iodine; RAI-Group-A, radiiodine with hyperthyroidism control by one-year, and RAI-Group-B, radiiodine without hyperthyroidism control by one-year; *Numbers add up to 171 for 170 patients because one patient had thyroidectomy and died in the first year.
**FIGURE 2: Kaplan Meier curves for mortality or MACE**

Kaplan Meier curves for cumulative mortality (panels A-D) or MACE (panels E-H) in the whole cohort and by primary treatment. ATD, patients who received antithyroid drugs alone in the first one-year, RAI-Group-A, radioiodine with hyperthyroidism control by one-year, and RAI-Group-B, radioiodine without hyperthyroidism control by one-year. HR, Hazard ratios derived from Cox regression models adjusted for age, sex, and comorbidity.
### FIGURE 3: Hazard ratios for mortality or MACE by primary treatment

HR (95% CI) are obtained from multivariable Cox proportional hazard regression models. ATD, patients who received antithyroid drugs alone in the first one-year, RAI-Group-A, radioiodine with hyperthyroidism control by one-year, and RAI-Group-B, radioiodine without hyperthyroidism control by one-year. P-value, for association; P-int, P values for interactions with comorbidity and age. No comorbidity=comorbidity score 0; comorbidity=comorbidity score >1
# FIGURE 4: Hazard ratios for mortality or MACE by one-year TSH

HR (95% CI) are obtained from multivariable Cox proportional hazard regression models. Analysis was conducted in all patients and in treatment subgroups according to treatment in the first-year with antithyroid drugs (ATD) or radioactive iodine (RAI). P-int, P values for interaction with treatment group.
FIGURE 5: One-year FT4 and mortality and MACE

Relationship between FT4 at one-year and All-cause mortality (A-C) and MACE (D-F). Curves represent log hazard ratios, 95% CI derived from restricted cubic spline regressions with 4-equally spaced knots at percentiles 5, 35, 65 and 95 respectively. FT4 is in multiples of the upper reference limit (x ULN) with the reference value set at FT4 x 0.75 ULN. Results were obtained from Cox regression models adjusted for age, gender, TRAb concentration, comorbidity, and year of diagnosis. P values are for overall association (P-overall) and non-linearity (P-non-linearity).