Incidental thyroid carcinoma in surgery-treated hyperthyroid patients with Graves’ disease: a systematic review and meta-analysis of cohort studies

Qingyi Jia,1,6* Xiaodan Li,2,6* Ying Liu,3,6 Ling Li,4 Joey SW Kwong,4,6 Kaiyun Ren,1 Yong Jiang,7 Xin Sun,4 Haoming Tian,1 Sheyu Li1
1Department of Endocrinology and Metabolism, West China Hospital, Sichuan University, Wuhou District, Chengdu, Sichuan, People’s Republic of China; 2Department of Gastroenterology, West China Hospital, Sichuan University, Wuhou District, Chengdu, Sichuan, People’s Republic of China; 3Department of Pediatrics, West China Second University Hospital, Sichuan University, Wuhou District, Chengdu, Sichuan, People’s Republic of China; 4Department of Pathology, West China Hospital, Sichuan University, Wuhou District, Chengdu, People’s Republic of China; 5Cochrane Taiwan, Taipei Medical University, Taipei, Taiwan; 6Department of Clinical Epidemiology and Department of Health Policy, National Center for Child Health and Development, Tokyo, Japan; 7Department of Pathology, West China Hospital, Sichuan University, Wuhou District, Chengdu, Sichuan, People’s Republic of China

*These authors contributed equally to this work

Introduction
The prevalence of thyroid carcinoma has rapidly increased over two-fold in the past three decades.1 The hypothetical mechanisms for the disease include exposure to ionizing radiation (especially among children), iodine deficiency and excess, etc.2 Hashimoto’s thyroiditis (HT), a common autoimmune thyroid disease,3–6 was reported to be associated with increased risk of incidental carcinoma.7,8 One explanation for this association was the underlying autoimmunity.9,10

Graves’ disease (GD), also known as toxic diffuse goiter, is another common autoimmune thyroid disease and is considered as the leading cause of hyperthyroidism.11 Although GD is not a traditional risk factor for thyroid cancer,13,14 a recent study15 showed that the expression levels of specific miRNAs in thyroid tissue from GD patients lay between those with normal thyroid glands and those with papillary thyroid carcinoma (PTC), suggesting the possibility of PTC development in GD patients. A population-based study also indicated a 10-fold increased cancer risk in GD patients.16 However, because the cancer detection by thyroid ultrasound...
examination and subsequent confirmation of diagnosis were more likely to be conducted in patients with GD compared to euthyroid adults, the conclusion defining GD as a risk factor of PTC cannot be drawn. Therefore, whether GD serves as a risk factor of thyroid carcinoma is still unclear. Three therapeutic options for hyperthyroidism were suggested by the guidelines, namely radio-iodine therapy, medication, and surgery. Incidental carcinoma could be identified occasionally in surgery-treated GD and non-GD hyperthyroid patients without identified malignancy before surgery. However, it remains unclear whether incidental thyroid carcinoma (ITC) is more frequently detected in patients with GD than those with other causes of hyperthyroidism. Thus, the current study aimed to investigate whether GD was associated with the risk of thyroid carcinoma in surgery-treated hyperthyroid patients by conducting a systematic review and comparing the prevalence of ITC between surgery-treated patients with GD and hyperthyroidism of other causes.

Materials and methods

Search strategy

This meta-analysis was planned, performed, and reported in accordance with the Meta-Analysis of Observational Studies in Epidemiology guidelines. Electronic databases were used for the literature searching, including EMBASE (from 1949) and PubMed (from 1946). The last search was updated to January 23, 2018. The combination of the keywords is given in the Supplementary materials. The Medical Subject Heading terminology was also used when available. As a supplement, bibliographies of identified studies were searched to make sure all the potentially relevant studies were included. No ethical approval or patient consent was required because all analyses were based on already existing literature.

Inclusion and exclusion criteria

Included studies had to have met the following inclusion criteria: 1) they should be cohort studies, either prospective or retrospective, 2) all patients in the studies underwent surgery due to hyperthyroidism, and 3) the number of ITC in both GD patients and non-GD hyperthyroid patients should be reported. Exclusion criteria were as follows: 1) duplicated studies or publications, 2) studies with incomplete data, 3) studies including any patient with diagnosed or suspected thyroid malignancy before surgery in the analysis, and 4) studies published in languages other than English.

Study selection

Two reviewers (QJ and YL) independently screened the titles and abstracts, and then full texts for eligibility. Discrepancies between the two reviewers’ selections were resolved by discussion with a third author (SL).

Data extraction and quality assessment

The following information was extracted by the three reviewers (QJ, XL, and SL) independently from every eligible study: first author, year of publication, country, ethnicity, surgery method, definition of GD and thyroid carcinoma, age, gender, patient-recruiting window, number of subjects with GD and non-GD hyperthyroidism (patients with toxic adenoma [TA], toxic nodular goiter [TNG]), and the number of thyroid carcinoma in these groups. If available, data on the histological diagnosis and proportion of microcarcinoma in GD and other thyroid diseases were collected. The Newcastle–Ottawa Scale was used for assessing the quality of the included studies. All disagreements were resolved through discussion.

Statistical analysis

In this meta-analysis, the prevalence was pooled by meta-analysis using STATA 12.0 (StataCorp LP, College Station, TX, USA). The risk of incidental thyroid cancer was measured by odds ratio (OR) with 95% confidence interval (CI), and pooled using Mantel–Haenszel (M–H) random-effect model due to clinical heterogeneity across studies. The overall data were pooled by meta-analysis, and a P-value below 0.05 was considered statistically significant. Between-study heterogeneity was estimated using the chi-squared test (in which $P<0.10$ was considered to indicate the presence of statistical heterogeneity) and quantified with the $I^2$ statistic (in which $I^2 \geq 50\%$ was considered to indicate significant statistical heterogeneity). Subgroup analyses by types of hyperthyroidism and microcarcinoma were also conducted to explore the source of heterogeneity. Five sensitivity analyses were carried out; these were as follows: 1) remove-one assay, 2) alternative pooling method (Peto instead of M–H method), 3) alternative statistical model (fixed effects instead of random effects), 4) exclusion of studies with ambiguous definitions of hyperthyroidism, and 5) exclusion of studies with unclear definitions of GD. Funnel graph, Begg’s test, and Egger’s test were employed to evaluate the potential publication bias. All statistical tests were performed using Review Manager 5.3 (Cochrane Collaboration, Oxford, UK) and STATA version 12.0 (STATA; StataCorp LP).
Results
Study screening and characteristics
As shown in Figure 1, a total of 8,416 articles were originally identified, and 74 articles were additionally identified through looking at the references of the screened papers and peer discussion. After title and abstract screening, 37 studies were left for full-text assessment. Eleven studies involving 10,743 GD patients and 3,336 patients without GD were included for quantitative analysis and were eligible for further meta-analysis. The baseline characteristics of each included study are summarized in Table 1. The mean age of patients from each included study ranged from 39 to 55.7 years. The majority (73–84.5%) of subjects were female. Seven studies had available data on thyroid microcarcinoma, which contributed to 65.59% of the total cancer occurrence. The definitions of GD and hyperthyroidism were based on clinical decisions and symptoms (diagnostic details are shown in Table S1). All definitions of thyroid cancer were based on postsurgical pathological results (histological definitions and classifications can be seen in Table S2). Besides GD, four studies included surgery-treated hyperthyroid patients with TA, while ten studies included those with TNG. The indications as well as types of thyroidectomy are also listed in Table S1.

Quality assessment
The Newcastle–Ottawa Scale of included studies ranged from 5 to 8, and the average score was 6.5 (Table 1). Three studies were scored at the lowest level (5 in score) mainly due to their failure to provide adequate information on whether all patients were recruited from the same population.

Meta-analysis
Among the included studies, the prevalence of ITC in surgery-treated hyperthyroid patients ranged from 0.4% to 32.4%, and the pooled prevalence was 7.0% (95% CI 4.5–9.6) (Figure 2). As shown in Figure 3, the prevalence of incidental carcinoma in GD and non-GD patients was not significantly different (pooled OR, 1.0; 95% CI: 0.68–1.46; P=0.98).

Figure 1 Flow diagram of study selection.
In the subgroup analysis, four studies involving 83 patients reported the ITC in patients with GD and TA, and ten studies compared the prevalence of ITC in GD and TNG (including toxic multinodular goiter and toxic uninodular goiter) patients, detecting no significant difference (pooled OR, 0.53; 95% CI: 0.21–1.36; \( P = 0.18 \) and pooled OR, 1.01; 95% CI: 0.65–1.57; \( P = 0.95 \), respectively).

PTC is the most common type of thyroid carcinoma, especially in incidental carcinoma.\(^{18}\) Eight studies\(^{28–31,33,34,36,37}\) reported incidental PTC with clear definition (including
single papillary carcinoma, occult papillary carcinoma, multifocal papillary carcinoma, and papillary microcarcinoma). The overall meta-analysis and subgroup analysis according to the etiology of hyperthyroidism provided similar results regarding the overall risk of ITC (Figure S1).

Publication bias and sensitivity analysis

Funnel plots seemed to be symmetric in our meta-analysis (Figure S2). The Egger's test ($P=0.774$) did not detect a statistically significant publication bias.

In order to assess the stability of the result of the current meta-analysis, sensitivity analysis was performed by sequentially excluding one single study each time. Similar results were acquired, suggesting the stability of our meta-analysis (Figure S3.1). As the definitions of either hyperthyroidism or GD were absent in several studies (Table S1), another sensitivity analysis was conducted by excluding studies lacking detailed explanation of hyperthyroidism or GD. The results showed no significant change after excluding undefined hyperthyroidism or GD (pooled OR, 1.16; 95% CI: 0.83–1.59; $P=0.38$, respectively; shown in Figures S3.2 and S3.3, respectively). Sensitivity analyses using Peto OR and M–H fixed-effect model also confirmed the robustness of the study (Figures S3.4 and S3.5).

Discussion

Our systematic review including 11 cohort studies and over 10,000 participants indicated that the pooled prevalence of ITC was 7.0% (95% CI: 4.5%–9.6%) in surgery-treated hyperthyroid patients. Unlike HT, GD did not seem to increase the risk of ITC compared with other hyperthyroid diseases among surgery-treated hyperthyroid patients. Both the subgroup analysis and sensitivity analysis confirmed the conclusion.

Although accumulated studies suggested an increased cancer risk in patients with HT and positive HT-related autoantibodies, our study demonstrated that no increase of ITC risk was detected in surgery-treated patients when compared to those with other causes. A possible explanation is that low or low–normal thyroid-stimulating hormone level could be present in GD patients before the thyrotoxicosis is diagnosed. Studies have proven that long-term suppressed level of thyroid-stimulating hormone (eg, in TA and TNG) significantly reduced the risk of thyroid carcinoma, especially PTC. This theory might explain the opposite risk of the thyroid malignancy in patients with HT and GD. Meanwhile, our study indicated that thyroid autoimmunity and presence of thyroid autoantibodies, which are common in both GD and HT, might contribute little in the pathogenesis of thyroid malignancy. But direct evidence is required in future investigations.

It must be noted that although ITC could be occasionally detected in hyperthyroid patients when their thyroid glands were removed by surgery, it should not be overtreated. The prevalence of occult papillary carcinoma in autopsy varied from 5.6% to 35.6%, indicating PTC as a latent neoplasm. Papillary thyroid microcarcinoma contributed to a very low mortality in the absence of treatment but with active surveillance. Since GD is not associated with the development of thyroid cancer according to our study, thyroidectomy should not be the first-line treatment in hyperthyroid patients where hyperthyroidism is caused by GD. Also, thyroidectomy should not be done for the purpose of preventing thyroid cancer.
Our subgroup analysis indicated that the risk of ITC was not significantly different among patients with hyperthyroidism caused by TNG/TA and GD. However, TNG and TA are both benign thyroid tumors, with a theoretical potential to turn malignant on occasion. Thus, choosing TNG and TA as the control in our study might have underestimated the malignancy risk of GD.

A recent population-based study in Taiwan indicated a 10-fold increased cancer risk in patients with GD, which was inconsistent with our results. The study was not included in our meta-analysis because the treatment modality of the participants was not identified. In this claims database analysis, the patients and controls were not well stratified by the risk of cancers during the matching and adjustment. Most of the controls seemed to be free of a history of thyroid diseases, compared to whom the GD patients in the study were more likely to undergo thyroid ultrasound examinations to find occult masses. A recent meta-analysis by Staniforth also reported on the ITC incidence among surgery-treated GD patients, revealing no increased risk of carcinoma. However, this study did not pool all cohort, case-control, and cross-sectional studies in a single analysis, which was considered methodologically inappropriate and might introduce additional bias. One study included in our analysis with a particularly large sample size was absent in theirs. Contributing to the major population in our meta-analysis, this study increased the strength of our study but did not lead to instable results, as can be seen by the sensitive analysis (Figure S3).

**Strengths and limitations of this study**

The strengths of our study include the comprehensive collection of current evidence of any study design for the topic and the nonselection of studies with patients with suspected malignancy. Some limitations of this meta-analysis must be acknowledged when interpreting our results. First, the definitions of presurgical GD and hyperthyroidism were mainly based on the clinical assessments and were heterogeneous between the included studies. But the multiple sensitivity analyses did not alter the trend of our results. Second, due to the lack of data of nodule presence in GD and the treatment length prior to surgery, related stratified analyses were not feasible. Whether the thyroid carcinoma was found in the nodule or the gland was not clarified among studies. It is difficult in this study to adjust thyroid nodule as a corisk factor of thyroid carcinoma. Meanwhile, although the statistical method did not indicate significant between-study heterogeneity, some clinical heterogeneity could still be undetected.

**Conclusion**

Our systematic review demonstrated that GD was not associated with increased risk of ITC in surgery-treated hyperthyroid patients. Thus, thyroidectomy is inappropriate for the purpose of preventing thyroid cancer in GD patients. Further long-term cohort study of cancer risk is required in hyperthyroid patients receiving nonsurgical treatment, such as antithyroid drugs or radioactive iodine therapy.

**Acknowledgments**

The study was supported by grants from the National Natural Science Foundation of China [grant number 81400811 and 21534008], National Basic Research Program of China [grant number 2015CB942800], the Scientific Research Project of Health and Family Planning Commission of Sichuan Province [grant number 130029, 150149, 17PJ063, and 17PJ445], Cholesterol Fund by China Cardiovascular Foundation and China Heart House, and the International Visiting Program for Excellent Young Scholars of SCU.

**Author contributions**

HT and SL conceived this study. QJ and YL performed the literature search and identification. KR assessed the full-text of included studies. QJ, XL, and SL extracted the data. QJ, LL, JSWK, and XS performed the statistical analysis. QJ, XL, YJ, HT, and SL interpreted the results. QJ, XL, KR, and SL drafted the manuscript. All authors critically revised the manuscript and agree to be accountable for all aspects of the work.

**Disclosure**

The authors report no conflicts of interest in this work.

**References**


