Use of Pre-Ablation Radioiodine-131 Scan to Assess the Impact of Surgical Volume and Specialisation following Thyroidectomy for Differentiated Thyroid Carcinoma

Beng Khiong Yap\textsuperscript{a}  Jarrod Homer\textsuperscript{b}  Neil Parrott\textsuperscript{b}  Sean Loughran\textsuperscript{b}  Brian Murby\textsuperscript{a}  Ric Swindell\textsuperscript{a}  Ashu Gandhi\textsuperscript{c}

\textsuperscript{a}The Christie NHS Foundation Trust, \textsuperscript{b}Central Manchester NHS Foundation Trust, and \textsuperscript{c}University of Manchester, Manchester Academic Health Science Centre, University Hospital of South Manchester NHS Foundation Trust, Manchester, UK

Key Words
Thyroid cancer · Surgeon volume · Thyroid remnant · Pre-ablation radioiodine scan · Surgical specialisation

Abstract

**Background:** We evaluated the relationship between thyroid remnant size following thyroidectomy for differentiated thyroid carcinoma and surgical volume and specialisation by assessing pre-ablation radioiodine-131 ($^{131}$I) thyroid bed uptake (TBU) scanning as a surrogate for residual thyroid tissue.

**Methods:** We analysed data of 651 patients in our thyroid cancer database. Patients' data were included if the following criteria were met: (1) diagnosis of differentiated thyroid carcinoma, (2) total or near-total thyroidectomy, (3) pre-ablation $^{131}$I scan prior to radioiodine ablation (RAI), (4) no distant metastasis, and (5) $>3,000$ MBq ablative dose of $^{131}$I. $^{131}$I diagnostic whole-body scans and measurement of thyroglobulin levels were carried out 3–9 months after RAI. 305 patients were included in the final analysis.

**Results:** Four endocrine, 19 otolaryngology and 25 general surgeons performed thyroidectomies with median pre-ablation $^{131}$I TBU values of 1.0, 1.8 and 2.9%, respectively (p = 0.0031). There was a statistically significant relationship between number of thyroidectomies performed and median pre-ablation $^{131}$I TBU values up to the optimal number of 11 operations beyond which there was no further significant difference between surgeons. There were differences in remnant size between endocrine and general surgeons (p = 0.001), otolaryngology and general surgeons (p = 0.023) but not between endocrine and otolaryngology surgeons (p = 0.167).

**Conclusion:** Using the pre-ablation $^{131}$I uptake scan as a surrogate for thyroid remnant quantification following thyroidectomy demonstrates the relationship between the surgical volume and size of thyroid remnant. The study also demonstrated beneficial effects of specialisation with specialist surgeons achieving the smallest thyroid remnant.

This work was presented in abstract form as a poster at the 82nd American Thyroid Association Annual Meeting, Quebec City, Quebec, Canada, September 19–23, 2012.
Introduction

The relationship between surgical volume and clinical outcome in endocrine surgery has been well reported in the literature [1–4]. The current American Thyroid Association and European Thyroid Association guidelines recommend total or near-total thyroidectomy for differentiated thyroid carcinoma (DTC) of >1 cm [5, 6]. Although the surgical mortality associated with thyroidectomy is low [2], the potential long-term complications of recurrent laryngeal nerve palsy and hypocalcaemia can be devastating. Thyroidectomy, in the hands of high-volume surgeons, has a less than 1% risk of recurrent laryngeal palsy [2, 7] as compared to 4.2% in low-volume surgeons [2]. The goals of initial therapy for DTC are to remove the primary tumour and thyroid tissue to facilitate post-operative remnant ablation with radioactive iodine-131 (RAI). The role of RAI is two-fold: firstly to ablate the residual normal thyroid remaining after total thyroidectomy, thus facilitating the early detection of recurrence based on serum thyroglobulin (Tg) measurement, and secondly it is to destroy persistent thyroid cancer cells after appropriate surgery as an adjuvant therapy employing the potential tumoricidal effect of RAI [5, 6].

Thyroidectomies are currently being performed by general, endocrine and otolaryngology (ENT) surgeons [2]. Increasingly, cancer surgeries are carried out by specialist surgeons as high surgeon operative volume has been shown to reduce surgical morbidity and mortality in oesophageal, lung and colorectal cancer resection [8–10]. The main benefit suggested for specialisation was that performing higher volumes of selective procedures would enhance surgical proficiency and in turn improve patient outcome. A systematic review by Chowdhury et al. [11] has demonstrated that high surgeon volume and specialisation are associated with improved patient outcome, while high hospital volume is of limited benefit. A recent publication examining the same issues in thyroid surgery concurs with these views [12].

The best clinical outcome of total thyroidectomy for DTC would be to achieve complete tumour resection and minimal or no thyroid remnant, thus avoiding the need for further surgery and radioiodine therapy. A pre-ablation radioiodine scan or measurement of $^{131}$I thyroid bed uptake (TBU) is useful when the extent of the thyroid remnant cannot be accurately ascertained from the surgical report, or when the results would alter either the decision to treat or the activity of RAI that is administered [5, 6]. We undertook a study to evaluate the effect of surgical specialisation on thyroid remnant size by measuring the radioiodine ($^{131}$I) uptake in the pre-ablation $^{131}$I scan.

Methods

Study Sample

We analysed the data of 651 patients in our institutional thyroid cancer database in 2004–2011. It was our institutional practice to perform a pre-ablation diagnostic $^{131}$I scan prior to RAI in 2004–2008. From 2009 onwards, there was a change in our practice, to perform only post-treatment $^{131}$I scans. Patients’ data were included if patients had (1) diagnosis of DTC, (2) total or near-total thyroidectomy, (3) pre-ablation $^{131}$I scan prior to RAI, (4) no distant metastasis, and (5) >3,000 MBq ablative dose of $^{131}$I. We identified a total of 305 patients who met the above criteria.

All patients were instructed to go on a low iodine diet for 3 weeks and discontinue levothyroxine ($T_4$) for 4 weeks or liothyronine ($T_3$) for 2 weeks prior to the diagnostic dose of $^{131}$I. The pre-ablation diagnostic $^{131}$I dose was 30–40 MBq and this was given 1 week prior to RAI. The diagnostic $^{131}$I scan was carried out 24 h following $^{131}$I administration. The pre-ablation diagnostic $^{131}$I scan was carried out using gamma cameras with high-energy collimators. The uptake value in the thyroid bed was corrected for soft-tissue background, and decay-corrected back to the time of administration.

One week after the diagnostic $^{131}$I administration, thyroid ablation was carried out using a target ablative $^{131}$I dose of 3,500 MBq ± 5%. Post-treatment $^{131}$I scans were not performed in this study group. Patients continued with thyroid hormone withdrawal and low iodine diet 4 days and 1 day after RAI respectively. Patients were started on $T_4$ 4 days after RAI.

Thyroid-stimulating hormone (TSH)-stimulated $^{131}$I diagnostic whole-body scans (DxWBS) and measurement of Tg and Tg antibody ($Tg$-Ab) levels either following thyroid hormone withdrawal or recombinant TSH (Thyrogen®) were carried out 3–4 months after RAI in 2004 and 6–9 months after RAI from 2005 onwards. Patients were also instructed to go on a low iodine diet 3 weeks prior to the diagnostic dose of $^{131}$I. The diagnostic $^{131}$I dose after RAI given was 150 MBq and DxWBS was performed 48 h after $^{131}$I administration.

Tg and Tg-Ab measurements were carried out using a radioimmunometric assay (Immulite). Our institutional cut-off for Tg was ≤1 ng/ml. The institutional cut-off for thyroid bed $^{131}$I uptake in DxWBS was ≤0.1% at 3–9 months following RAI. The sixth edition of the TNM staging system was used for this study.

Statistical Analysis

We classified the number of cases of thyroidectomies that a particular surgeon had carried out according to all the 651 records in our database. We assessed thyroid cancer surgery experience by counting how many operations they did out of the 651 cases. This could mean for some surgeons that in their earlier cases, if they started practicing near to 2004, the surgeon was actually inexperienced counting how many operations they did out of the 651 cases. This could mean for some surgeons that in their earlier cases, if they started practicing near to 2004, the surgeon was actually inexperienced although because they did many operations later they are defined as experienced for the purposes of this study. If a patient underwent ipsilateral hemithyroidectomy followed by contralateral thyroidectomy (two-stage thyroidectomies) it would be classed as 1 case.

We split the 305 cases by the ‘thyroid cancer surgery experience’ of their surgeons by comparing the median pre-ablation $^{131}$I TBU values starting with the cases where the surgeon had done only 1 case versus those where the surgeon did ≥2 cases. We then moved the ‘experience’ bar to cases where surgeons had done 1–2 versus ≥3 cases and so on up to 1–12 versus ≥13 cases in a stepwise manner. We applied the Mann-Whitney U test to compare the median TBU
values between the two ‘surgical’ group cases in each of the above. We used the p value from each test to determine the optimal number of cases required to show that the median TBU values between the low- and high-volume surgeons is no longer statistically significant.

We also grouped the surgeons into three categories: general, endocrine and ENT surgeons. Endocrine surgeons were those whose operative workload was 40% or more of endocrine surgery. General surgeons were those who performed regular thyroid surgery but this represented <40% of their operative workload. ENT surgeons were those where training and operative practice was exclusively confined to ear, nose and neck surgery including thyroid operations. The pre-ablation 131I thyroid bed median uptake values among the three different categories of surgeon were analysed using Kruskal-Wallis one-way non-parametric analysis of variance test.

If the Kruskal-Wallis test was significant, then, in order to assess the effect of specialisation in the three different categories of surgeons, we carried out further analyses to determine the difference in the pre-ablation 131I thyroid bed median uptake values in the categorized group ENT versus general, general versus endocrine and ENT versus endocrine using the Mann-Whitney U test for each paired comparison by applying a Bonferroni correction and so used a significance value α = 0.05/3 = 0.017.

The relationship between the pre-ablation diagnostic 131I TBU values and the TSH-stimulated DxWBS TBU values and Tg levels were analysed using the χ² test. Statistical analysis was performed with SPSS software version 16.0 (SPSS, Inc.). Statistical significance was defined as p ≤ 0.05 unless otherwise specified.

**Results**

**Study Group**

305 patients met the study criteria and their data were analysed. The characteristics of the entire cohort of patients including TNM staging are summarised in table 1. The details of RAI and DxWBS for the study group are summarised in table 2. The pre-RAI TSH levels were available in 241 subjects and the median TSH was 99.36 mU/l (range 6.12 to >150). The 131I DxWBS were performed at median 6.3 months (range 2.8–9.7). This was carried out following thyroid hormone withdrawal (THW) in 218 patients, whilst 87 patients had recombinant TSH (Thyrogen™) and stayed on T₄.

**Surgeons and Surgical Volume**

A total of 48 surgeons performed thyroidectomies in 651 patients. Out of 651 patients, 305 had pre-ablation whole-body diagnostic 131I scan. We used data of all 651 patients to analyse surgical volume and the available (305 cases) median pre-ablation 131I TBU value to examine the relationship between surgical volume and remnant size. This in our view would provide a more meaningful representation of any potential relationship between volume and remnant size.

<table>
<thead>
<tr>
<th>Table 1. Characteristics of the 305 patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Median age, years</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2. RAI and 131I DxWBS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median dose 131I given, MBq (range)</td>
</tr>
<tr>
<td>Method of TSH stimulation</td>
</tr>
<tr>
<td>Patients, n</td>
</tr>
<tr>
<td>Median TSH levels, mU/l</td>
</tr>
<tr>
<td>131I DxWBS, months</td>
</tr>
<tr>
<td>Patients, n</td>
</tr>
</tbody>
</table>

There were 25 general, 19 ENT and 4 endocrine surgeons. Table 3 summarises the category of surgeon and the total number of cases performed as well as the median pre-ablation 131I TBU value associated with different types of surgeon. The median pre-ablation 131I TBU values for endocrine, ENT and general surgeons were 1.0, 1.8 and 2.9%, respectively. The overall effect of specialisation was statistically significant as demonstrated by p = 0.0031 with the endocrine surgeons achieving the lowest median 131I scan uptake values. Further analysis using Bonferroni correction (where significance level α ≤ 0.017) showed a significant difference between general and endocrine surgeons (p = 0.001) and a trend towards between ENT and general surgeons (p = 0.023) but no significant difference between ENT and endocrine surgeons (p = 0.167).

The relationship between surgical volume and median pre-ablation 131I TBU values is summarised in table 4. There was a statistically significant difference in the median uptake values starting with the group of surgeons who had only performed 1–4 cases versus surgeons who...
had done ≥ 5 cases. This difference continued up to cases where their surgeons had done between 1 and 10 cases versus the group of surgeons had done ≥ 11 cases. The p value then no longer showed statistical significance for cases where surgeons performed more than 11 cases.

**Tg and Pre- and Post-Ablation Diagnostic \(^{131}\)I Uptake Scans**

278 patients had pre-ablation \(^{131}\)I TBU values of >0.1% and 27 had an uptake value of ≤ 0.1% as shown in table 5. A total of 289 patients underwent DxWBS at 3–9 months,
with 80 patients (27.7%) having the DxWBS at <6 months and 209 patients (72.7%) having the DxWBS ≥6 months following radioiodine ablation. Seven out of 283 patients had >0.1% of $^{131}$I uptake in the thyroid bed 3–9 months following RAI. Therefore, the successful ablation rate based on DxWBS was 97.5%. Only 247 out of 305 patients had Tg and Tg-Ab measurement at the time of their DxWBS 3–9 months post-RAI. 25 patients had detectable Tg-Ab. The number of patients with Tg level ≤1 and >1 ng/ml were 184 and 38, respectively. Therefore, the successful ablation rate based on Tg measurement was 83% excluding patients who had Tg-Ab as the antibodies interfered with Tg assay.

There was no statistically significant association between the pre-ablation $^{131}$I TBU value and the TBU value in the DxWBS 3–9 months post-RAI ($p = 0.51$). Similarly, there was no statistically significant association between the pre-ablation $^{131}$I TBU values and the Tg measurement at the time of their DxWBS 3–9 months post-RAI ($p = 0.51$). We also analysed the relationship between different levels of pre-ablation $^{131}$I TBU values (0–0.1, 0.2–10 and >10%) with the Tg measurement at the time of their DxWBS 3–9 months post-RAI, which again showed no statistically significant difference ($p = 0.287$).

**Discussion**

Our study was carried out to assess the impact of surgical volume and specialisation of individual surgeons and the quantity of remnant thyroid tissue following thyroidectomy for DTC, as measured by pre-ablation $^{131}$I scan TBU value, over a period of 5 years (2004–2009). The factors influencing the amount of thyroid remnant are likely to be complex depending on the peri-operative findings relating to tumour extent, surgical experience and training.

This study was not designed to examine surgical complication rates related to surgical volume as this has been extensively published previously [1–4, 7]. Sosa et al. [1] reported a study of 5,860 thyroidectomies carried out over a 6-year period, concluding that the highest-volume surgeons (>100 cases) were most likely to operate on patients with thyroid cancer, and that these had the lowest complication rates and shortest lengths of hospital stay. Similarly, Stavakis et al. [3] reported that surgeons performing ≥100 cases per year maintain the lowest complication rates.

A national audit conducted by the British Association of Endocrine and Thyroid Surgeons (BAETS) examined the data of 2,500 thyroid surgical cases submitted between 2005 and 2009 [2]. The audit showed that the rates of temporary hypocalcaemia stratified by groups <11, 11–25, 26–50 and >50 cases per year were 17.6, 27.2, 34.1 and 13.5%, respectively. The rate of recurrent laryngeal nerve palsy stratified by groups <11, 11–25, 26–50, >50 cases per year were 4.2, 3.1, 4.0 and 0.9%, respectively. The conclusions from the audit were that surgeons performing >50 cases per year have the lowest complication rates.

Our data show that the median pre-ablation $^{131}$I TBU values in the group of surgeons performing 1–11 versus >11 cases onwards were not statistically significant (table 4). This would suggest that the optimal number of cases of DTC operated by the study surgeons’ group was >11. Using the cut-off of 11 cases, the optimal number of cases a thyroid surgeon who carries out benign and malignant thyroid surgeries needs to perform to achieve the lowest pre-ablation $^{131}$I TBU values can be estimated. The BAETS audit showed that thyroid cancer surgery accounts for approximately 20% of thyroid surgical workload in the UK. Therefore, based on our estimation, the optimal total number of thyroid cases (benign and malignant) would be 55 annually to achieve the lowest amount of residual thyroid tissue (as measured by pre-ablation $^{131}$I TBU values). Our data was based on the cases in our database over a period of 5 years. Therefore, from our study, we are unable to recommend a specific number of thyroidectomies per annum required to achieve the smallest thyroid remnant. Nevertheless, our findings would further strengthen the conclusion of the BAETS audit where >50 thyroid surgical cases per year are required to achieve the lowest complication rates. Although there is no consistent definition of what constitutes ‘satisfactory’ surgeon volume in thyroid surgery, our study findings add to the body of evidence supporting the association between surgeon volume and clinical outcome [1–4, 7, 13].

The systematic review by Chowdhury et al. [11] demonstrated the independent contributions of specialisation and volume of surgery. When the volume discrepancies between specialist and general surgeons were adjusted, the benefit from specialisation was still apparent. There are various factors that can influence the differences in the clinical outcomes between specialist and general surgeons, including the cumulative effect of high volume, more effective choices of treatment modalities and factors unrelated to surgeons such as contributions from other healthcare professionals and the effect of selective referral.
Our study demonstrated a statistically significant difference in the median pre-ablation $^{131}$I TBU values between specialist and general surgeons as shown in table 3. The median pre-ablation $^{131}$I TBU values for the endocrine, ENT and general surgeons were 1, 1.8 and 2.9%, respectively ($p = 0.0031$). When the Bonferroni correction was applied, the significance level for $\alpha \leq 0.017$, the effect of specialisation was apparent between the endocrine and general surgeons ($p = 0.001$) and a trend towards significance between the ENT and general surgeons ($p = 0.023$). This observation may be due to the influence of case mix between high- and low-volume ENT surgeons. There was no statistically significant difference between endocrine and ENT surgeons. Current European guidelines for the treatment of DTC state that total/near-total thyroidectomy facilitates post-surgical RAI and adequate follow-up and that if there is evidence of more limited thyroidectomy further completion thyroidectomy surgery may need to be considered [6]. Our data shows potential benefits in surgical specialisation in achieving minimal thyroid remnants following thyroidectomy and near or total thyroidectomy and this remains the recommended strategy for DTC in accordance with European guideline.

There are limitations in our study: First, we do not have the data to differentiate whether the operations were carried out by trained surgeons or surgeons in training which could influence the results. However, the BAETS audit showed that thyroid cancer surgery was more likely to be carried out by the consultant (trained surgeon) in the UK [2]. Second, we cannot be certain that all thyroid cancer cases were referred, although our institution is the only regional centre with the facility for radioiodine therapy. Third, our statistical analysis does not take into account the potential improvement in surgical skills when surgeons become more experienced with time. Fourth, our database only collects data for thyroid cancers. However, we assessed thyroid cancer surgery experience by grouping surgeons based how many cases they did out ever, we assessed thyroid cancer surgery experience by grouping surgeons based how many cases they did out. Lastly, the study did not set out to examine the surgical complications as a measure of clinical outcome. Furthermore, as this is a retrospective study, we do not know what the impact of leaving a large thyroid remnant actually has on recurrence. However, historical data showed recurrence was more common in patients treated with subtotal thyroidectomy in DTC.

Our data would suggest that the radioiodine activity given in this study was effective, even in patients with large thyroid remnants as measured by the pre-ablation diagnostic $^{131}$I scan. This finding was in keeping with the observation reported by Leblanc et al. [14] where similar recurrence rates were found in patients who were treated with total thyroidectomies only, or partial thyroidectomies plus RAI. The lack of statistical association between pre-ablation $^{131}$I thyroid remnant assessment and Tg levels and DxWBS would suggest that pre-ablation $^{131}$I diagnostic scan is not necessary if standard 3,700 MBq is used as recommended by the American Thyroid Association and European Thyroid Association [5, 6]. There were two different timings of diagnostic WBS 3–4 and 6–9 months post-RAI due to a change in institutional practice and 6 out of 305 patients did not have diagnostic WBS due to various reasons including declining further investigations and lost to follow-up. However, the results of the diagnostic WBS post-RAI have no impact on the conclusion of the study. Furthermore, the successful ablation rate in our study was comparable to the published literature [15], suggesting that pre-ablation $^{131}$I diagnostic scan was not detrimental to clinical outcome. With the publication of randomised evidence from the British and French studies [16, 17] supporting the use of low activity (1,100 MBq) of radioiodine in DTC, the use of pre-ablation $^{131}$I diagnostic scanning may still be of value where the extent of the thyroid remnant cannot be accurately ascertained, especially when thyroidectomies are performed by non-specialist surgeons, and when the results would alter either the decision to treat or the activity of RAI that is administered [18]. Although this study did not examine the impact of thyroid remnant size and clinical outcome, future study of this nature should be pursued to reaffirm the near or total thyroidectomy strategy in the treatment of DTC.

In summary, we have found that using the pre-ablation $^{131}$I TBU scan as a surrogate for quantification of thyroid remnant following thyroidectomy for DTC demonstrates the relationship between the volume of surgery and size of thyroid remnant. Our data also suggest a beneficial effect of specialisation with specialist surgeons achieving minimal thyroid remnants represented by the lowest pre-ablation $^{131}$I TBU values.

**Disclosure Statement**

The authors have no conflicts of interest to disclose.
References


