

# Difficulties in deciding whether to ablate patients with putatively “low–intermediate-risk” differentiated thyroid carcinoma: do guidelines mainly apply in the centres that produce them? Results of a retrospective, two-centre quality assurance study

Savvas Frangos<sup>1</sup> · Ioannis P. Iakovou<sup>2</sup> · Robert J. Marlowe<sup>3</sup> · Nicolaos Eftychiou<sup>1</sup> · Loukia Patsali<sup>1</sup> · Anna Vanezi<sup>1</sup> · Androulla Savva<sup>1</sup> · Vassilis Mpalaris<sup>2</sup> · Evanthia I. Giannoula<sup>2</sup>

Received: 5 March 2015 / Accepted: 29 June 2015 / Published online: 1 August 2015  
© Springer-Verlag Berlin Heidelberg 2015

## Abstract

**Purpose** We determined the reasons for radioiodine thyroid remnant ablation, and the procedure’s necessity based on post-surgical remnant size, in patients with putatively “low–intermediate-risk” differentiated thyroid carcinoma (DTC). We identified key clinicopathological, treatment and remnant characteristics, and factors associated with remnant size in 336 patients with pT1/2, M0 DTC ablated during the period September 2010 to October 2013 at one Cypriot or one Greek referral centre.

**Methods** Clinicopathological/treatment characteristics were compiled from charts. Experienced nuclear medicine physicians rated the numbers/intensities of uptake foci in the thyroid bed on postablation planar scintigrams using scales of 0–4 points and 0–3 points, respectively. The product of these scores was taken as the “remnant score” that

ranged from 0 (no remnant) to 12 (multiple remnants, intense uptake).

**Results** DTC was predominantly papillary. The median [25th–75th percentile] longest primary tumour diameter was 1.0 cm [0.7–1.5 cm]. Despite favourable histotypes and primary tumour classifications, patients often had preablation characteristics suggesting elevated or uncertain risk: 31.0 % of patients (104 of 336) had primary tumour multifocality, 22.0 % (74) had confirmed cervical lymph node metastases, 37.2 % (125) had unknown nodal status, and 38.1 % (128) had antithyroglobulin antibody seropositivity. The median [25th–75th percentile] remnant score was 4 [2–6]; 39.9 % of patients (134 of 336) had scores  $\geq 6$ . For the entire cohort, T or N stages ( $r \leq 0.174$ ,  $P \leq 0.05$ ) correlated positively with the remnant score in a univariate Spearman analysis. The numbers of patients referred by the surgeon, cervical lymph nodes excised and metastatic nodes excised correlated negatively ( $r \leq 0.243$ ,  $P \leq 0.038$ ) with the remnant score, and the first two factors independently predicted the remnant score ( $P \leq 0.037$ ) in a multivariate analysis.

**Conclusion** Patients with putatively “low–intermediate-risk” DTC frequently had disease characteristics denoting high or uncertain risk, suggesting that “selective” radioiodine ablation in such patients may seldom be applicable outside international centres of excellence. Proxies for surgeon experience and surgical completeness correlated with remnant number/uptake intensity and may aid ablation-related decision-making.

Some of the data in this paper were delivered in an abstract and an oral presentation at the 37th Annual Meeting of the European Thyroid Association, Leiden, the Netherlands, 7–11 September 2013.

**Electronic supplementary material** The online version of this article (doi:10.1007/s00259-015-3124-4) contains supplementary material, which is available to authorized users.

✉ Savvas Frangos  
savvas.frangos@gmail.com

<sup>1</sup> Department of Nuclear Medicine, Bank of Cyprus Oncology Centre, 32 Acropoleos Avenue, Strovolos, Nicosia 2006, Cyprus

<sup>2</sup> Department of Nuclear Medicine, Aristotle University, Papageorgiou Hospital, 56403 Thessaloniki, Greece

<sup>3</sup> Spencer-Fontayne Corporation, 33 Bentley Avenue, Jersey City, NJ 07304-1901, USA

**Keywords** Low-risk DTC · Thyroid remnant ablation · Clinicopathological · Thyroid remnant size · Guidelines · Surgeon experience

## Introduction

Postoperative thyroid remnant ablation with “therapeutic” activities of  $^{131}\text{I}$  has traditionally been applied in most (near-)totally thyroidectomized patients with differentiated thyroid carcinoma (DTC). Three main rationales have been offered for this practice [1–3]. First, ablative radioiodine administration allows highly sensitive posttherapy whole-body scintigraphy (rxWBS) to discover previously undetected disease. Second, the ablative activity also treats microscopic or other occult radioiodine-sensitive tumour, possibly decreasing DTC recurrence and mortality [4, 5]. Last, by eliminating sources of thyroglobulin (Tg) secretion and radioiodine uptake, ablation should improve the sensitivity of serial monitoring for recurrence using serum Tg testing or diagnostic whole-body scintigraphy (dxWBS).

However, these benefits, especially benefits related to outcome, can be difficult to prove. Most published data in the DTC setting are retrospective and observational. Additionally, studies may be underpowered or of insufficient duration to detect infrequent, slow-to-emerge events, e.g. recurrence or cause-specific death, in patients with this generally relatively indolent neoplasm with high survival [6, 7]. Thus, and because radioiodine therapy sometimes has side effects [8–10], recent DTC management guidelines (for example references [11–15]) generally advocate “selective” use of this procedure in patients considered to be at “low-risk” or even “intermediate-risk” of DTC recurrence or mortality. Thus the American Thyroid Association 2009 DTC management guidelines [11] include patients with T1–2 or N1 disease in the “selective ablation use” category, and indeed, recommend against ablation for N0–Nx disease. T1–2 and N1 patients also should be among the “selective use” population according to the National Comprehensive Cancer Network DTC management guidelines [15]. However, although they are valuable compendia of state-of-the-art data and disease management suggestions, DTC management guidelines are usually formulated by physicians from international centres of excellence, and may not be fully applicable elsewhere [16, 17]. For example, clinicopathological characteristics themselves, or their documentation, may be less clear-cut in other centres, and higher-risk features such as anti-Tg autoantibodies (TgAb) may be more prevalent in certain countries or regions.

Despite guideline recommendations, at our centres, we have in recent years continued to use radioiodine ablation in nearly all patients with a primary tumour  $\leq 4$  cm and no evidence of distant metastasis, i.e. pT1–2 M0 status. We make these therapeutic decisions on an individualized, patient-by-patient basis in everyday practice. In our centres, such decision-making generally does not rely on post-surgical neck ultrasonography or scintigraphic imaging; the former may be difficult to interpret due to postoperative inflammation [11] and the latter may impose excessive

logistical demands on patients, may slow down workflow, or may require prohibitively expensive isotopes (e.g.  $^{123}\text{I}$ ) to avoid “stunning”.

For these reasons, it can be challenging to gain a “big picture” of the factors leading to the decision to ablate in this putatively “low–intermediate-risk” population. We therefore performed the present retrospective quality assurance study with three objectives. First, we sought to determine key preablation clinicopathological and treatment characteristics of pT1–2 M0 DTC patients ablated in our centres during a recent period of about 3 years, i.e. the factors influencing our decision to ablate. Second, through retrospective evaluation of postablation scintigrams regarding the number and intensity of foci of thyroid bed uptake, we aimed to assess thyroid remnant size in this cohort. This endpoint reflects the necessity for ablation, at least provided that rationales for the procedure are accepted. Our third goal was to identify clinicopathological or treatment characteristics associated with thyroid remnant size before ablation. Such variables might help identify “surgically-ablated” patients, and hence could help improve ablation-related decision-making.

## Materials and methods

### Patients, setting, and ethics

The study included 336 patients with classical or follicular variants of papillary DTC or with follicular DTC (including Hürthle cell histology), and with pT1–2 M0 status before ablation. This cohort comprised all such patients with available medical records who from 17 September 2010 to 31 October 2013 underwent a first radioiodine ablation procedure at the Nuclear Medicine Departments of Bank of Cyprus Oncology Centre (BOCOC), Strovolos Nicosia, Cyprus ( $n=198$ ) or Papageorgiou Hospital (PGH), Thessaloniki, Greece ( $n=138$ ). These institutions are the largest tertiary referral centres for DTC in Cyprus and Northern Greece, respectively. The cohort comprised >99 % and about 95 %, respectively, of patients with pT1–2 M0 papillary or follicular DTC initially presenting at BOCOC or PGH during the study period. The patients seen during this period who were not included in this study (nine patients for both centres combined) had missing medical records (two patients), refused ablation due to fears regarding radiation (six patients) or refused consent for use of their data in scientific studies/publications (one patient).

Ethics committee approval for this study was not sought, since the analysis was anonymous and retrospective, and therefore neither affected patient privacy nor entailed additional interventions. Moreover, before radioiodine treatment, patients provided written informed consent allowing use of their data in such analyses.

## Surgery and ablation

Patients had to have undergone (near-)total thyroidectomy. Ablation took place a median of 73 days [interquartile range, IQR, 52–136 days] after surgery (334 patients). Patients were prescribed a low-iodine diet for the 10 days before the procedure.

All but one patient at BOCOC and 64.5 % at PGH (89/138) had ablation stimulated by thyroid hormone withdrawal (THW; no thyroid hormone for the 2–6 weeks before ablation). One patient at BOCOC and 35.5 % at PGH (49/138) had ablation stimulated by the approved regimen of recombinant thyroid-stimulating hormone (rhTSH, Thyrogen; Genzyme, Cambridge, MA) comprising two consecutive daily intramuscular injections of 0.9 mg. Within a few days after surgery, the patients receiving rhTSH were placed on thyroid hormone therapy, which was not interrupted or altered for ablation.

TSH at ablation was <30 mIU/L in 0.6 % of patients with available data (2 of 334), but was always >20 mIU/L. Patients were ablated with fixed empirical radioiodine activities, typically (280 of 336 cases, 83.3 %) 3.7 GBq (100 mCi). Patients were hospitalized for 2–4 days after ablation, and were discharged when the exposure rate at 1 m was <40  $\mu$ Sv/h, unless discharge was medically contraindicated.

## rxWBS

rxWBS was performed 5–10 days after ablative radioiodine administration. Anterior and posterior planar images of the cranium, thorax and abdomen from the top of the head to the inguinal region, and spot images as needed, were obtained while the patients were supine. Large field-of-view double-headed gamma cameras (BOCOC: Infinia Hawkeye 4GP3, GE Healthcare, Tirat Carmel, Israel; PGH: ADAC dual-head, Philips, Amsterdam, The Netherlands) equipped with high-energy collimators were used. Scanning was performed for  $\geq$ 500,000 counts (about 30 min). To minimize workflow disruption and time and logistical demands on patients, neck uptake was not quantified.

## Remnant size scoring

Thyroid remnant was retrospectively given classification scores based on visual assessment of the number of foci of thyroid bed uptake and their overall intensity on postablation scintigrams. The classification scores were derived as the product of two subscores, the number of thyroid bed foci and the overall uptake intensity. The number of thyroid bed foci was objectively scored as: 0 no foci, 1 one focus, 2 two foci, 3 three foci, and 4 more than three foci. The overall uptake intensity was subjectively scored, without reference to other tissues, e.g. liver, as: 0 no uptake, 1 low uptake, 2 intermediate uptake, and 3 star effect. (Representative

postablation WBS images illustrating scores for intensity of uptake and maximum remnant scores are shown in Fig. 1.) Thus there were nine possible remnant size scores, ranging from 0, no apparent remnant, to 12, multiple remnants, high uptake.

Remnant classification scoring for each centre's images was performed by two experienced nuclear medicine physicians at the respective institutions (N.E. and S.F. at BOCOC; E.I.G. and I.P.I. at PGH). Each physician pair worked jointly; disagreements were resolved by consensus between one scorer from each centre (S.F. and I.P.I.). The scorers from the respective centre were involved in the patients' treatment; however, images were labelled only with five-digit patient identification numbers, and were evaluated approximately 1–3 years after ablation, without reference to other patient data.

## Histopathological classification

Histopathological classification was performed by the pathologist working with the given surgeon, based on the surgeon's report as well as the pathologist's analysis of excised tissue. The histology report and the American Joint Committee on Cancer/Union Internationale contre le Cancer Tumour, Nodes, Metastasis (TNM) system, 6th edition [18] were used to stage each patient's disease before ablation. Patients were classified as free of cervical lymph node metastasis (N0) if one or more such nodes were excised and all excised nodes were negative for DTC on pathological examination. Patients not meeting these criteria but without evidence of cervical node involvement before ablation were classified as Nx.

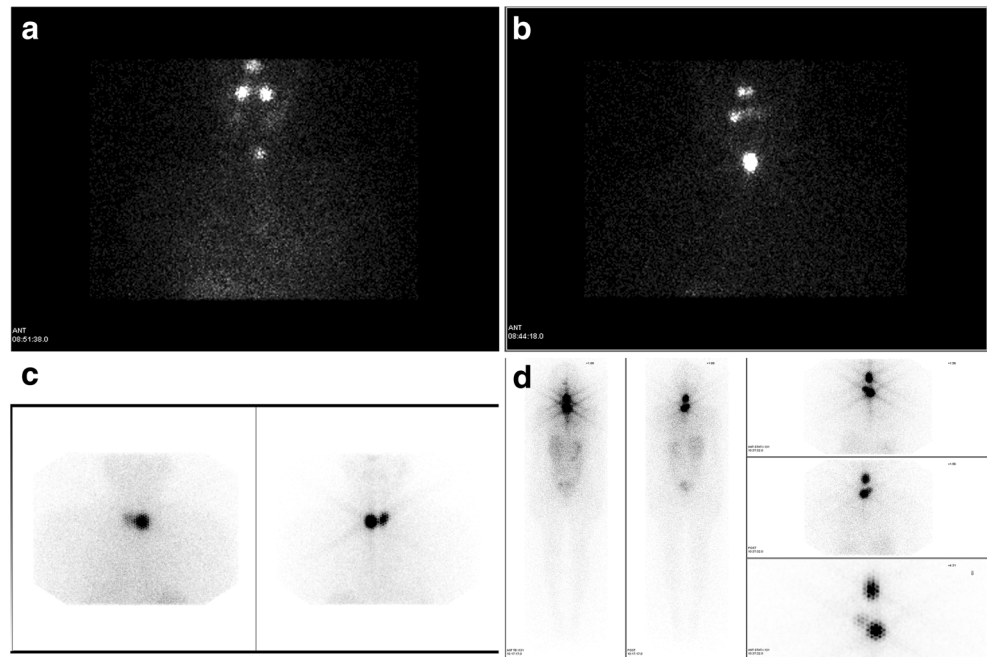
## Biochemistry

Tg and TgAb were measured using commercial assays (see Supplementary Table 1 for methodological details) by one accredited central laboratory for each institution (an independent firm at BOCOC; part of the Nuclear Medicine Department at PGH). Throughout the study period, each laboratory used a single assay for each of these two analytes. Samples were drawn immediately before ablative radioiodine administration in the 286 patients undergoing THW or 72 h after the second rhTSH injection, i.e. 48 h after ablative administration, in the 50 patients receiving rhTSH.

## Statistics

Discrete variables are expressed as counts and percentages, or vice versa, continuous variables as median (minimum–maximum) or median [IQR]. Patient characteristics were compared between centres using the chi-square test or Mann-Whitney test, as appropriate.

**Fig. 1** Representative postablation WBS images illustrating scores for intensity of uptake and maximum remnant scores: **a** score 1, low uptake, **b** score 2, intermediate uptake, **c** score 3, star effect, **d** maximum remnant score of 12 (foci subscore of 4, intensity subscore of 3)



Spearman's rank order correlation coefficient was assessed to explore the strength and direction of associations between predefined variables of clinical interest and the thyroid remnant score as a surrogate for remnant size. The variables comprised age at ablation, DTC histology, number of primary tumours, T stage, primary tumour multifocality, maximum primary tumour diameter, N stage, TSH stimulation method, surgeon's referral score, number of cervical lymph nodes excised, and number of metastatic cervical nodes excised. The referral score was assigned on a scale of 1–4 points based on the number of patients operated on by a given surgeon who were referred to BOCOC or PGH during the study period. The score comprised a proxy for the surgeon's experience in thyroid surgery. Surgeons referring <5 patients received a referral score of 1; 5–9 patients, a score of 2; 10–19 patients, a score of 3; and  $\geq 20$  patients, a score of 4. Age, maximum primary tumour diameter, number of cervical lymph nodes excised, number of excised cervical lymph nodes with DTC involvement, and referral and remnant scores were analysed as continuous variables. Analysed as categorical variables were: DTC histology (follicular vs. papillary, classical variant vs. papillary, follicular variant, vs. Hürthle cell), T stage (T1 vs. T2), multifocality (yes vs. no), N stage (N0 vs. N1 vs. Nx), TSH stimulation method (THW vs. rhTSH).

Two-tailed  $P$  values of the Spearman correlations were calculated based on a test statistic assuming a  $t$ -distribution with 334 degrees of freedom. Independent variables with a statistically significant Spearman correlation with the remnant score were then included in a multiple linear regression model, with the dependent variable remnant score as a continuous variable. A stepwise backward elimination procedure was applied by

eliminating one at a time the factors with the highest  $P$  values  $\geq 0.05$ .

Testing was two-tailed;  $P < 0.05$  was considered to be statistically significant. SPSS version 18.0.1 (IBM SPSS, Armonk, NY) was used.

## Results

### Cohort characteristics: comparison between centres

Table 1 summarizes and compares key preablation characteristics, and Table 2, Tg and TgAb findings, by study centre. At both centres, roughly three-quarters of the patients were female and tended to be in early middle age.

The cohorts from the two centres differed significantly regarding DTC histology, T and N status before ablation, median longest diameter of the primary tumour, capsule infiltration, surgeons' referral score, cervical lymph nodes excised before ablation, and rates of Tg <1  $\mu\text{g/L}$ , TgAb positivity, and unavailable data regarding TgAb. Specifically, the BOCOC patients overwhelmingly comprised patients with the classical papillary histotype, whereas fewer than 40 % of the PGH patients had this characteristic – half had follicular variant papillary thyroid cancer. Virtually all the BOCOC patients had T1 disease, whilst slightly over one-quarter of the PGH patients had T2 disease. In line with these T status profiles, the median longest diameter of the primary tumour was 50 % larger in the PGH patients than in their BOCOC counterparts. However, the PGH patients tended to have less advanced and more frequently

**Table 1** Patient characteristics

Characteristic	BOCOC ( <i>n</i> =198)	PGH ( <i>n</i> =138)	<i>P</i> value, BOCOC vs. PGH
Female, % ( <i>n</i> )	79.3 (157)	72.5 (100)	0.153
Age at ablation			
Median years [IQR]	43.0 [34.0–53.0]	44.5 [34.3–52.0]	0.695
≥45 years, % ( <i>n</i> )	43.9 (87)	50.0 (69)	
DTC histology, % ( <i>n</i> )			<b>&lt;0.001</b>
Papillary, classical variant	94.9 (188)	37.7 (52)	
Papillary, follicular variant	5.1 (10)	50.0 (69)	
Follicular	0 (0)	10.9 (15)	
Hürthle cell	0 (0)	1.4 (2)	
T status <sup>a</sup> at ablation, % ( <i>n</i> )			<b>&lt;0.001</b>
Any T1	90.4 (179)	73.9 (102)	
T1a	61.6 (122)	38.4 (53)	
T1b	28.8 (57)	35.5 (49)	
Any T2	9.6 (19)	26.1 (36)	
Primary tumours	( <i>n</i> =195)	( <i>n</i> =137)	
Median number [IQR]	1 [1–2]	1 [1–2]	0.575
Multifocal, % ( <i>n</i> )	29.3 (58)	33.3 (46)	0.550
Longest diameter of primary tumour	( <i>n</i> =197)		<b>&lt;0.001</b>
Median (cm) [IQR]	0.8 [0.6–1.3]	1.2 [0.9–2.2]	
>1 cm, % ( <i>n</i> )	34.3 (68)	61.6 (85)	
Capsule infiltration, % ( <i>n</i> )			<b>&lt;0.01</b>
None	50.0 (99)	60.9 (84)	
Thyroid infiltration	8.6 (17)	23.2 (32)	
Nodule infiltration	26.8 (53)	15.9 (22)	
No data available	14.6 (29)	0 (0)	
N status <sup>a</sup> at ablation, % ( <i>n</i> )			<b>&lt;0.001</b>
N0	49.5 (98)	28.3 (39)	
Nx	25.8 (51)	53.6 (74)	
Any N1	24.7 (49)	18.1 (25)	
Number of patients referred per surgeon, median [IQR]	1 [1–3.75]	2 [1–4]	0.433
Surgeons' referral score <sup>b</sup> , median [IQR]	1 [1–1]	1 [1–1]	0.435
Surgeons' referral score category, % ( <i>n</i> )			<b>&lt;0.001</b>
1	80 (35)	85 (39)	
2	11 (5)	11 (5)	
3	5 (2)	4 (2)	
4	5 (2)	0 (0)	
Cervical lymph nodes excised			<b>&lt;0.001</b>
Median number [IQR]	4 [1–9]	0 [0–3]	
None, % ( <i>n</i> )	24.2 (48)	54.3 (75)	
1 or 2, % ( <i>n</i> )	4.0 (8)	18.1 (25)	
≥10, % ( <i>n</i> )	24.7 (49)	8.0 (11)	
Excised cervical lymph nodes with DTC involvement			0.101
Median [IQR]	0 [0–3]	0 [0–0]	
None, % ( <i>n</i> )	74.2 (147)	81.9 (113)	

Percentages may not add exactly to 100 % due to rounding. Patient characteristics were compared between centres using the chi-square test or Mann–Whitney test, as appropriate. *P* values considered to be significant (<0.05) are in bold type

BOCOC Bank of Cyprus Oncology Centre, PGH Papageorgiou Hospital

<sup>a</sup> TNM staging was according to the American Joint Committee on Cancer/Union Internationale contre le Cancer staging system, 6th edition [18]

<sup>b</sup> Based on the number of patients that the surgeon operated on and referred to BOCOC or PGH during the study period, the surgeon's referral score ranged from 1 to 4, with higher scores denoting more patients referred



**Table 2** Tg and TgAb findings

Characteristic	BOCOC ( <i>n</i> =198)	PGH ( <i>n</i> =138)	<i>P</i> , BOCOC vs. PGH
Evaluable <sup>a</sup> stimulated Tg, % ( <i>n</i> )	57.6 (114)	62.3 (86)	0.206
Stimulated Tg category, % ( <i>n</i> )			
<1 µg/L	10.1 (20)	4.0 (8)	<b>0.005</b>
1–<5 µg/L	19.7 (39)	30.3 (60)	0.458
5–10 µg/L	8.1 (16)	7.3 (6)	0.585
>10 µg/L	19.7 (39)	12.3 (12)	0.755
TgAb-positive, % ( <i>n</i> )	41.9 (83)	32.6 (45)	<b>&lt;0.001</b>
No TgAb data, % ( <i>n</i> )	0.5 (1)	5.1 (7)	<b>&lt;0.01</b>

Patient characteristics were compared between centres using the chi-square test or Mann-Whitney test, as appropriate. *P* values considered to be significant (<0.05) are in bold type

BOCOC Bank of Cyprus Oncology Centre, PGH Papageorgiou Hospital

<sup>a</sup>Patients were considered to have evaluable Tg values if they were TgAb-negative on direct testing

characterized capsule infiltration than did the BOCOC patients. Perhaps reflecting their larger median primary tumour size, PGH patients significantly less frequently had Tg <1 µg/L than did the BOCOC patients. TgAb positivity was significantly more frequent among the BOCOC patients than among the PGH patients, although it was frequent among patients from both centres. Virtually all BOCOC patients had available data regarding TgAb status, but about one in twenty PGH patients lacked such data.

The PGH thyroid surgeons operated on fewer patients than did the BOCOC surgeons; the PGH surgeons excised fewer cervical lymph nodes and less frequently removed such nodes. Related to the latter, the PGH cohort included a far larger percentage of patients with Nx disease than did the BOCOC cohort.

### Cohort characteristics and risk classification

Despite favourable histotypes and primary tumour classifications, substantial proportions of patients at both centres had characteristics that suggested higher risk or uncertainty regarding risk, or both. Thus about one-third of each cohort had multifocal primary tumour. The largest tumour diameter in the thyroid was ≥1 cm in more than one-third of the BOCOC patients and in more than three-fifths of the PGH patients. Almost one-quarter of the BOCOC patients and almost one-fifth of the PGH patients had known cervical lymph node metastases before ablation. Reflecting frequently limited cervical lymph node excision, another quarter of BOCOC patients and a majority of PGH patients had unknown cervical lymph node status before ablation.

Regarding biochemistry, at ablation, more than two-fifths of BOCOC patients and almost one-third of PGH patients were TgAb-positive. Of TgAb-negative patients with available Tg measurements, 82.5 % of BOCOC patients (94/114), 90.7 % of PGH patients (78/86), and 86.0 % overall (172/200) had Tg concentrations ≥1 µg/L.

Figure 2 shows the overall pattern of patient referrals by surgeons at BOCOC and PGH. Altogether 90 surgeons performed thyroid surgery on the 336 patients. Only six of these surgeons (7 %) operated on ten or more patients, and 41 (46 %) operated on only a single patient.

### Thyroid remnant size

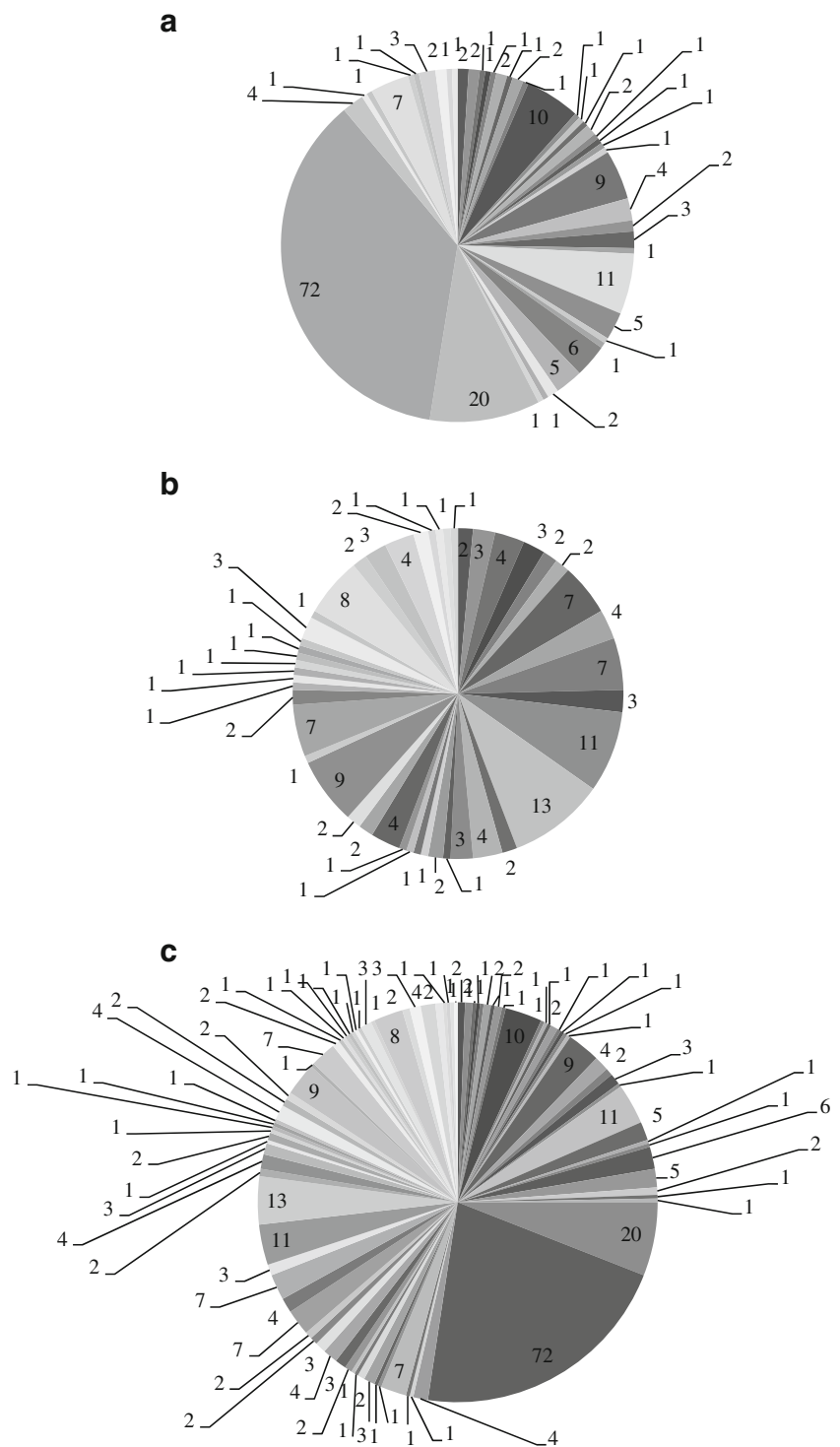
Table 3 presents data regarding thyroid remnant size, based on retrospective assessment of the number and intensity of foci of thyroid bed uptake on postablation planar scintigrams. Virtually all patients had evidence of gross postsurgical thyroid remnant (thyroid bed uptake was absent in only 4 % of the BOCOC patients and in 1.4 % of the PGH patients). Remnant was substantial in almost one-third of BOCOC patients and almost one-half of PGH patients, as reflected by remnant scores of 6, 8, 9 and 12, the four highest of nine possible scores.

### Factors associated with thyroid remnant scores

Table 4 shows Spearman correlations between pre-defined factors of clinical interest and thyroid remnant scores for the overall cohort and by centre. Also presented are the results of a multivariate analysis of the relationships of these variables.

In the univariate analysis, for the entire cohort and each centre's cohort, the surgeon's referral score and the number of cervical lymph nodes excised were significantly negatively correlated with the remnant score. T stage was significantly correlated with the remnant score for the entire cohort, but the relationship did not attain significance for either centre alone. N stage was significantly correlated with the remnant score for the entire cohort and the BOCOC cohort, but not the PGH cohort. Conversely, the number of metastatic cervical nodes excised was significantly negatively correlated with the remnant score for the entire cohort and the PGH cohort, but not the BOCOC cohort.

**Fig. 2** Numbers of patients operated on by each surgeon at **a** BOCOC, **b** PGH, and **c** both centres combined. At BOCOC, altogether 44 surgeons performed thyroid surgery on 198 patients; at PGH, altogether 46 surgeons performed thyroid surgery on 138 patients. At BOCOC, only 4 of the 44 surgeons (9 %) and at PGH, only 7 of the 46 surgeons (15 %) each operated on >5 % of patients



Backward multiple regression analysis revealed that the dependent variable, remnant score, could be independently predicted by the surgeon’s referral score for the entire cohort ( $P=0.025$ ) and for the BOCOC cohort ( $P=0.017$ ), but not for the PGH cohort ( $P=0.307$ ). The number of cervical lymph nodes excised independently predicted the remnant score for the entire cohort ( $P=0.037$ ) and the PGH cohort ( $P=0.005$ ),

but not the BOCOC cohort ( $P=0.35$ ). T stage, N stage, and number of excised metastatic lymph nodes did not significantly predict remnant score for the entire cohort or for the individual centre cohorts for which these variables were tested. These findings suggest that surgeon experience or the completeness of neck surgery predicts postsurgical, preablation thyroid remnant size.

**Table 3** Thyroid remnant scores

Score	Cohort, % (n)		
	All patients (N=336)	BOCOC (n=198)	PGH (n=138)
0	3.0 (10)	4.0 (8)	1.4 (2)
1	11.9 (40)	13.1 (26)	10.1 (14)
2	17.9 (60)	19.7 (39)	15.2 (21)
3	5.7 (19)	4.0 (8)	8.0 (11)
4	21.7 (73)	26.3 (52)	15.2 (21)
6	18.2 (61)	18.7 (37)	17.4 (24)
8	9.5 (32)	7.6 (15)	12.3 (17)
9	4.8 (16)	3.0 (6)	7.2 (10)
12	7.4 (25)	3.5 (7)	13.0 (18)
Median score [IQR]	4 [2–6]	4 [2–6]	4 [2–6]

BOCOC Bank of Cyprus Oncology Centre, PGH Papageorgiou Hospital

Remnant scores were determined from posttherapy planar scintigrams by two experienced nuclear medicine physicians from the respective study centres, and were the product of a four-point score for the number of foci and a qualitative three-point score for the intensity of thyroid bed uptake. There were nine possible remnant classification scores, ranging from 0 (no apparent remnant) to 12 (multiple remnants, high uptake)

**Table 4** Results of Spearman and multivariate regression analyses of factors potentially related to thyroid remnant score

Variable	Spearman correlation						Backward multiple regression		
	All		BOCOC		PGH		All	BOCOC	PGH
	Coefficient (r)	P value <sup>a</sup>	Coefficient (r)	P value	Coefficient (r)	P value	P value	P value	P value
Age at ablation	0.056	0.30	0.032	0.65	0.060	0.41	NIM	NIM	NIM
DTC histology	0.240	0.66	0.020	0.69	0.070	0.36	NIM	NIM	NIM
Number of primary tumours	0.005	0.92	0.070	0.28	0.090	0.26	NIM	NIM	NIM
<b>T stage</b>	0.107	<b>0.05</b>	0.110	0.98	0.035	0.68	0.207	NIM	NIM
Primary tumour multifocality	0.006	0.91	0.060	0.34	0.080	0.34	NIM	NIM	NIM
Maximum primary tumour diameter	0.085	0.12	0.059	0.40	0.010	0.90	NIM	NIM	NIM
<b>N stage</b>	0.174	<b>0.001</b>	0.162	0.02	0.107	0.21	0.091	0.124	NIM
TSH stimulation method	0.400	0.46	0.050	0.42	0.040	0.58	NIM	NIM	NIM
<b>Surgeon's referral score<sup>b</sup></b>	-0.187	<b>0.001</b>	-0.182	<b>0.01</b>	-0.086	<b>0.01</b>	<b>0.025</b>	<b>0.017</b>	0.307
<b>Number of cervical lymph nodes excised</b>	-0.243	<b>&lt;0.001</b>	-0.159	<b>0.02</b>	-0.203	<b>0.017</b>	<b>0.037</b>	0.35	<b>0.005</b>
<b>Number of excised cervical lymph nodes with DTC involvement</b>	-0.113	<b>0.038</b>	-0.060	0.37	-0.150	<b>0.05</b>	0.571	NIM	0.743

Remnant scores were determined from posttherapy planar scintigrams by two experienced nuclear medicine physicians from the respective study centres, and were the product of a four-point score for the number of foci and a qualitative three-point score for the intensity of thyroid bed uptake. There were nine possible remnant classification scores, ranging from 0 (no apparent remnant) to 12 (multiple remnants, high uptake)

Variables in bold type are positively or negatively correlated with the remnant score ( $P < 0.05$ ). These variables were included as independent variables in a multiple linear regression model that had the dependent variable of remnant score as a continuous variable. A stepwise backward elimination procedure was then applied by eliminating factors with the largest  $P$  values ( $\geq 0.05$ ) one at a time

DTC differentiated thyroid cancer, NIM not included in model, rhTSH recombinant human thyroid-stimulating hormone, TSH thyroid-stimulating hormone, THW thyroid hormone withdrawal

<sup>a</sup> Two-tailed  $P$  values of the Spearman correlations were calculated based on a test statistic assuming a  $t$  distribution with 334 degrees of freedom

<sup>b</sup> Based on the number of patients operated on and referred to BOCOC or PGH by a surgeon during the study period, the surgeon's referral score ranged from 1 to 4, with higher scores denoting more patients referred



## Discussion

This retrospective analysis of numerous patients with putatively “low–intermediate-risk” pT1–2 M0 DTC ablated over a recent period of about 3 years at two tertiary referral centres had three main findings. First, virtually all patients had at least one clinicopathological characteristic suggesting higher risk of recurrence, disease-specific mortality, or both, or creating uncertainty as to risk stratification. Thus according to three leading recent consensus statements, 61.3–80.6 % of our cohort would have had probable or definite indications for ablation (Table 5).

Second, based on the number and intensity of thyroid bed uptake foci on postablation scans, virtually all (97 %) our patients had gross thyroid remnant after surgery, and a considerable proportion (39.8 %) had a remnant score  $\geq 6$ , i.e. one of the four highest of nine possible remnant size scores. This finding suggests that the decision to ablate was appropriate in most of our patients, if one accepts the desirability of eliminating thyroid remnant to improve the sensitivity of surveillance for recurrence using Tg testing, dxWBS, or both, or to remove sites of potential polyclonal malignant transformation [19]. Notably, even among the ten patients with a remnant score of 0, indicating no visible thyroid bed uptake on postablation scintigrams, six were TgAb-positive.

Third, in multivariate analyses, two surgery-related factors were significantly independently associated with thyroid remnant size for the entire cohort and for one of the individual centre cohorts. The first of these factors, the referral score, was based on the number of patients referred for ablation per surgeon. This variable presumably reflected surgeon experience in thyroid operations, at least in regions such as ours where few tertiary referral centres for DTC exist and surgeons routinely refer DTC patients for nuclear medicine evaluation. The second of these factor, the number of cervical lymph nodes dissected, presumably reflected the thoroughness of DTC excision. Both surgeon experience and completeness of surgery are well-documented as critical to patient outcome [20–25]. Indeed, our observations support the concept of radioiodine ablation as a “backstop” when an appreciable possibility exists of sub-optimal surgery [13]. It should be noted, however, that both

factors identified as significantly related to remnant size for the entire cohort were also found to be significant only for one of the two centre cohorts, and did not approach  $P=0.05$  for the other cohort. These observations may raise a caveat regarding the generalizability of the multivariate findings.

The prevalence of certain clinicopathological and data characteristics among our patients merits further comment. The high frequency of serum Tg  $\geq 1$   $\mu\text{g/L}$  (observed in 172 of 200 evaluable patients, 86.0 %, i.e. TgAb-negative patients with available Tg data) is unsurprising given the appreciable remnant volume in most of our cohort. The high rate of TgAb positivity (38.1 %, 128 of 328 patients with TgAb data) may also reflect the presence of gross remnant. Additionally, given that these autoantibodies have a half-life of about 10 weeks [26] and the comparable median time between the thyroid operation and TgAb measurement in our patients, TgAb positivity in some patients may have been a transient response to surgically stimulated antigen release into the circulation. However, persistent TgAb positivity increases reliance on dxWBS for follow-up [27], and hence the desirability of remnant ablation.

The present quality assurance survey clearly revealed frequent limitations in surgical treatment in our cohort and deficiencies in histopathological reports regarding many patients. A total of 90 surgeons operated on our patients. Most referred only modest numbers of patients (median 2 [IQR 1–4] for the entire cohort), often creating uncertainty about the surgical protocol and the completeness of resection. A further source of doubt about the completeness of resection, as well as the true extent of disease, was the frequently non-existent or very limited cervical lymph node dissection (no nodes excised in 123 of the 336 patients, 36.6 %, only one or two nodes removed in another 33 patients, 9.8 %). Additionally, due to our surgical referral pattern, our patients’ histopathological reports were from many different institutions, depended on input from surgeons of diverse experience levels, and took many forms. Classifications regarding capsular invasion appeared to be particularly inexact. These observations regarding surgical referrals and histopathology reports suggest that programmes to improve DTC patient care in our regions would be valuable. Such programmes might include efforts to educate generalist surgeons and pathologists regarding DTC, to direct referrals

**Table 5** Classification of the study population in relation to the indications for thyroid remnant ablation according to selected DTC clinical guidelines/consensus statements

Category/criterion	EANM 2008 [13]	NCCN version 2 2014 [15]	ATA 2006 [11]
Definite ablation candidates (“indication”)	70.5 % (237)	19.6 % (66)	32.7 % (110)
Probable ablation candidates (“recommendation”)	10.1 % (34)	54.8 % (184)	28.6 % (96)
Possible (“selective”) ablation candidates (“no indication”)	19.3 % (65)	25.6 % (86)	38.7 % (130)

The data presented are percentages (number) of patients in relation to the entire cohort

ATA American Thyroid Association, EANM European Association of Nuclear Medicine, NCCN National Comprehensive Cancer Network

for thyroid surgery to a smaller number of centres, and to standardize and ensure thoroughness of histopathology report contents.

Some limitations of our study should be considered. First, there were appreciable imprecision, heterogeneity and gaps in the data. For example, thyroid remnant size was assessed using an unvalidated, partly subjective scoring system. Our remnant score relied in part on quantification of foci of uptake. Since larger remnant foci with high uptake can effectively obscure further smaller foci, patients with large remnants may have had artificially low remnant scores. This phenomenon might have introduced bias by resulting in conservative remnant scoring. Nonetheless, our remnant size rating relied on the highly sensitive modality of rxWBS, and was conducted by four experienced nuclear medicine physicians. Additionally, our remnant size scoring system was chosen to reflect “real-life” conditions in our countries, where centres typically do not perform uptake measurement, SPECT/CT, or  $^{123}\text{I}$  scintigraphy. Additionally, Tg and TgAb results might have been affected by variations in the time since surgery [IQR 52–136 days] and in TSH stimulation method (rhTSH in 50 of the 336 patients, 14.9 %). It should be noted, however, that treatment guidelines [14, 15] basing indications for ablation on Tg or TgAb levels make no distinctions regarding indications based on time since surgery or TSH preparation.

Another gap in our data was ultrasonography results. However, in the initial weeks after surgery, oedema and other signs of inflammation may confound sonogram interpretation. Moreover, a disadvantage of neck ultrasonography is that its accuracy is highly operator-dependent. Most importantly regarding heterogeneity, imprecision, and gaps in our data, they are inherent to a retrospective study. However, these limitations have the virtue of reflecting real-world conditions and real-world inputs to decision-making regarding whether to ablate or not to ablate, the latter a main endpoint of our study.

It should be noted that the present investigation did not address outcomes or harm related to thyroid remnant ablation, including effects on patient quality of life, DTC recurrence, cause-specific or overall survival, or side effects. Additionally, the study took place at tertiary referral centres in the European Union. Our results may not be generalizable to other geographic areas, and may be conservative with respect to less specialized settings or more resource-constrained regions. Further, our cohort included only pT1–2 M0 DTC patients referred for nuclear medicine evaluation, and this that might have biased patient inclusion towards those more likely to have indications for ablation. However, to our knowledge, our local surgeons refer essentially all pT1–2 M0 patients for nuclear medicine evaluation, suggesting that any such bias was unlikely to have materially affected our findings.

In conclusion, our large two-centre retrospective quality assurance study found that patients with putatively “low–intermediate-risk” DTC frequently had higher-risk features, or

characteristics confounding risk stratification. This finding suggests that outside international centres of excellence, limitations in surgical experience and completeness and in histopathology reporting may cast important doubt on the classification of such patients as “low-risk” or “intermediate-risk”.

We also noted that our patients often had considerable thyroid remnant despite putative (near-)total thyroidectomy. Our observations suggest that “selective use” of radioiodine ablation even in pT1–2 M0 DTC patients may seldom be feasible outside international centres of excellence; this hypothesis is in line with the finding of a recent systematic review that applicability may be the greatest weakness of current DTC treatment guidelines [17]. However, surgeon experience, as reflected by the number of patients referred for possible ablation, and completeness of surgery, as reflected by the number of cervical lymph nodes excised, may aid in choosing ablation candidates.

**Acknowledgment** The support in the form of a grant from Genzyme, a Sanofi company, is acknowledged.

#### Compliance with ethical standards

**Funding** This was an investigator-initiated study funded from the regular budgets of the Departments of Nuclear Medicine of the two participating centres. However, the work of one of the authors (R.J.M.) in developing this paper was supported by a grant from Genzyme, a Sanofi company, the rhTSH manufacturer.

**Conflict of interest** The work of one of the authors (R.J.M.) in developing this paper, including descriptive statistical analysis, organization and presentation of data, drafting and editing the manuscript, and handling part of manuscript submission logistics, was supported by a grant from Genzyme, a Sanofi company, the rhTSH manufacturer. This author is an independent medical editor, whose clients include Genzyme. The other authors declare that they have no conflicts of interest.

**Ethical approval** This was a retrospective study, entailing no additional procedures. However, all procedures described here were in accordance with the ethical standards of the institutional and/or national research committees and with the principles of the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards.

**Informed consent** For this type of study, formal consent is not required. However, written informed consent was obtained from all patients included in the study before they underwent the radioiodine treatment described here. This consent included allowing the use of their data in analyses such as those presented here.

#### References

- Schlumberger MJ. Papillary and follicular thyroid carcinoma. *New Engl J Med.* 1998;338:297–306.
- Verburg FA, Dietlein M, Lassmann M, Luster M, Reiners C. Why radioiodine remnant ablation is right for most patients with differentiated thyroid carcinoma. *Eur J Nucl Med Mol Imaging.* 2009;36:343–6. doi:10.1007/s00259-008-0997-5.

3. Mazzaferri EL, Jhiang SM. Long-term impact of initial surgical and medical therapy on papillary and follicular thyroid cancer. *Am J Med.* 1994;97:418–28.
4. Tuttle RM, Lopez N, Leboeuf R, Minkowitz SM, Grewal R, Brokhin M, et al. Radioactive iodine administered for thyroid remnant ablation following recombinant human thyroid stimulating hormone preparation also has an important adjuvant therapy function. *Thyroid.* 2010;20:257–63.
5. Sawka AM, Brierley JD, Tsang RW, Thabane L, Rotstein L, Gafni A, et al. An updated systematic review and commentary examining the effectiveness of radioactive iodine remnant ablation in well-differentiated thyroid cancer. *Endocrinol Metab Clin North Am.* 2008;37:457–80.
6. McLeod DS, Sawka AM, Cooper DS. Controversies in primary treatment of low-risk papillary thyroid cancer. *Lancet.* 2013;381:1046–57.
7. Nixon IJ, Ganly I, Patel SG, Palmer FL, Di Lorenzo MM, Grewal RK, et al. The results of selective use of radioactive iodine on survival and on recurrence in the management of papillary thyroid cancer, based on Memorial Sloan-Kettering Cancer Center risk group stratification. *Thyroid.* 2013;23:683–94.
8. Grewal RK, Larson SM, Pentlow CE, Pentlow KS, Gonen M, Qualey R, et al. Salivary gland side effects commonly develop several weeks after initial radioactive iodine ablation. *J Nucl Med.* 2009;50:1605–10.
9. Almeida JP, Sanabria AE, Lima EN, Kowalski LP. Late side effects of radioactive iodine on salivary gland function in patients with thyroid cancer. *Head Neck.* 2011;33:686–90.
10. Molinaro E, Leboeuf R, Shue B, Martorella AJ, Fleisher M, Larson S, et al. Mild decreases in white blood cell and platelet counts are present one year after radioactive iodine remnant ablation. *Thyroid.* 2009;19:1035–41.
11. Cooper DS, Doherty GM, Haugen BR, Kloos RT, Lee SL, Mandel SJ, et al. Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer. *Thyroid.* 2009;19:1167–214.
12. Pacini F, Schlumberger M, Dralle H, Elisei R, Smit JW, Wiersinga W. European consensus for the management of patients with differentiated thyroid carcinoma of the follicular epithelium. *Eur J Endocrinol.* 2006;154:787–803.
13. Luster M, Clarke SE, Dietlein M, Lassmann M, Lind P, Oyen WJ, et al. Guidelines for radioiodine therapy of differentiated thyroid cancer. *Eur J Nucl Med Mol Imaging.* 2008;35:1941–59.
14. Silberstein EB, Alavi A, Balon HR, Clarke SE, Divgi C, Gelfand MJ, et al. The SNMMI practice guideline for therapy of thyroid disease with <sup>131</sup>I 3.0. *J Nucl Med.* 2012;53:1633–51.
15. Tuttle RM, Haddad RI, Ball DW, Dickson P, Duh Q-Y, Ehya H, et al. NCCN clinical practice guidelines in oncology: thyroid carcinoma. Version 2.2014. National Comprehensive Cancer Network; 2014. p. 1–120.
16. Dietlein M, Verburg FA, Luster M, Reiners C, Pitoia F, Schicha H. One should not just read what one believes: the nearly irresolvable issue of producing truly objective, evidence-based guidelines for the management of differentiated thyroid cancer. *Eur J Nucl Med Mol Imaging.* 2011;38:793–8.
17. Huang TW, Lai JH, Wu MY, Chen SL, Wu CH, Tam KW. Systematic review of clinical practice guidelines in the diagnosis and management of thyroid nodules and cancer. *BMC Med.* 2013;11:191.
18. Green F, Page DL, Fleming ID, Fritz AG, Balch CM, Haller DG, et al. *AJCC Cancer staging manual.* 6th ed. Chicago: Springer; 2002.
19. Shattuck TM, Westra WH, Ladenson PW, Arnold A. Independent clonal origins of distinct tumor foci in multifocal papillary thyroid carcinoma. *New Engl J Med.* 2005;352:2406–12.
20. Pisello F, Geraci G, Lo Nigro C, Li Volsi F, Modica G, Sciume C. Neck node dissection in thyroid cancer. A review. *G Chir.* 2010;31:112–8.
21. Schneider DF, Mazeh H, Chen H, Sippel RS. Lymph node ratio predicts recurrence in papillary thyroid cancer. *Oncologist.* 2013;18:157–62.
22. Schneider DF, Sippel RS. Measuring quality in thyroid cancer surgery. *Adv Endocrinol.* 2014;2014, 714291. doi:10.1155/2014/714291
23. Sosa JA, Bowman HM, Tielsch JM, Powe NR, Gordon TA, Udelsman R. The importance of surgeon experience for clinical and economic outcomes from thyroidectomy. *Ann Surg.* 1998;228:320–30.
24. Bilimoria KY, Bentrem DJ, Ko CY, Stewart AK, Winchester DP, Talamonti MS, et al. Extent of surgery affects survival for papillary thyroid cancer. *Ann Surg.* 2007;246:375–81. Discussion 381–4.
25. Bilimoria KY, Zanoocco K, Sturgeon C. Impact of surgical treatment on outcomes for papillary thyroid cancer. *Adv Surg.* 2008;42:1–12.
26. Spencer C. Commentary on: implications of thyroglobulin antibody positivity in patients with differentiated thyroid cancer: a clinical position statement. *Thyroid.* 2013;23:1190–2.
27. Verburg FA, Luster M, Cupini C, Chiovato L, Duntas L, Elisei R, et al. Implications of thyroglobulin antibody positivity in patients with differentiated thyroid cancer: a clinical position statement. *Thyroid.* 2013;23:1211–25.