Surgery 165 (2019) 608–616

Contents lists available at ScienceDirect

Surgery

journal homepage: www.elsevier.com/locate/surg

Endocrine

Comparison of short-term oncologic outcome of robotic thyroid surgery using dynamic risk stratification: A propensity score–matched comparison study

Jae Won Cho, MD[a], Jae Hyun Park, MD, PhD[b], Yu-mi Lee, MD, PhD[a], Suck Joon Hong, MD, PhD[a], Jong Ho Yoon, MD, PhD[b,*]

[a] Department of Surgery, Asan Medical Center, University of Ulsan College of Medicine, Wonju, Gangwon-do, Korea
[b] Department of Surgery, Wonju Severance Christian Hospital, Yonsei University Wonju College of Medicine, Wonju, KorGangwon-do, Korea

ARTICLE INFO

Article history:
Accepted 31 July 2018
Available online 13 September 2018

ABSTRACT

Background: The long-term oncologic outcome of robotic surgery for thyroid cancer is not well established. The aim of this study was to predict the long-term oncologic outcome of robotic surgery by using dynamic risk stratification in classic papillary thyroid carcinoma.

Methods: A total of 444 propensity score–matched pairs of patients with papillary thyroid carcinoma treated with robotic surgery and conventional open surgery were classified into 4 response-to-therapy categories. The results were compared between the robotic surgery and open surgery groups.

Results: The median follow-up duration was 60 months. After propensity score matching, the robotic surgery group showed less extensive thyroid surgery and lymph node dissection and a higher proportion of patients who underwent radioactive iodine remnant ablation than the open surgery group; however, the dynamic risk stratification did not differ between the 2 groups ($P = .086$).

Conclusion: The long-term oncologic outcome of robotic surgery is expected to be comparable with that of open surgery based on the dynamic risk stratification.

Introduction

Robotic thyroid surgery is being performed for properly selected patients with differentiated thyroid cancer, and its inclusion criteria are gradually expanding based on surgeon experience. Although many earlier published studies have validated robotic thyroid surgery with respect to technical safety, quality of life, and surgical completeness, the long-term oncologic outcome of robotic thyroid surgery for thyroid cancer has not been clearly established thus far because the median follow-up period of patients treated with robotic thyroid surgery has not yet reached 10 years.[1–9]

Dynamic risk stratification (DRS), based on the response to initial therapy, assessed 2 years postoperatively or at any time during follow-up, was validated as having a better capability in predicting long-term structural recurrence or persistent structural disease than other staging or recurrence risk stratification systems, especially in patients treated with total thyroidectomy (TT) and subsequent radioactive iodine (RAI) remnant ablation.[10–13] Furthermore, DRS is widely used to reclassify initial risk estimates based on clinicopathologic characteristics available just after initial therapy even in patients treated without RAI remnant ablation, with modification of the definitions of each response-to-therapy category.[14,15] Reportedly, the prediction of variance explained (a statistical tool for identifying how much the system can predict the target outcome) that values of the DRS in patients treated with TT and RAI remnant ablation ranged 62%–84%, which is higher than those of the American Thyroid Association (ATA) risk stratification system (25%–34%).[16–18] Furthermore, each response-to-therapy category has been shown to be significantly correlated with the development of structural recurrence or recurrence-free survival in the patient cohort with a median follow-up of 7–10 years.[12–15,18–21]

The aims of this study were to evaluate the DRS in patients who underwent robotic surgery (RS) for classic papillary thyroid carcinoma (PTC) and compare it with that of patients who underwent conventional open surgery (OS), and to deduce the long-term oncologic outcome of robotic thyroid surgery.

* Corresponding author: Department of Surgery, Wonju Severance Christian Hospital, Yonsei University Wonju College of Medicine, 20 Ilsan-ro, Wonju, Gangwon-do 26426, Korea.
E-mail address: gsyoon@yonsei.ac.kr (J.H. Yoon).

https://doi.org/10.1016/j.surg.2018.07.038
0039-6060/

Materials and Methods

Study population

Between December 2008 and April 2014, 2,349 consecutive patients underwent an initial operation (495 RS and 1,854 OS) for classic PTC, performed by a single endocrine surgeon at Asan Medical Center (Seoul, Korea). After excluding patients who underwent lateral lymph node (LN) dissection or had distant metastasis, those with a postoperative follow-up period <2 years, or those with an insufficient medical record, 2112 patients (492 and 1,620 treated with RS and OS, respectively) were enrolled in this study. Data were obtained from the prospectively maintained endocrine surgery database at Asan Medical Center (Seoul, Korea). The study protocol was approved by our institutional review board; the requirement for informed consent from each patient was waived because of the noninterventional nature of the study.

In this analysis, the extent of extrathyroidal extension (ETE) was divided into 3 categories, defined according to the 7th edition of the American Joint Committee on Cancer/Union for Cancer Control tumor-node-metastasis (TNM) staging system, as follows: "none," for intrathyroidal lesion without ETE; "minimal," for T3 tumors with microscopic extension; and "extensive," for T4 tumors.[22] Thereafter, the first 2 categories were combined to create a "nonextensive" category.

For the enrolled study participants, the clinicopathologic characteristics, including DRS, were evaluated and compared between the RS and OS groups. After excluding 78 patients with unavailable data for assessing response to therapy, based on the identifiable preoperative clinicopathologic parameters (age, sex, primary tumor size, ETE, multifocality, and bilaterality) that could affect the choice of the surgical procedure (RS or OS), propensity scores were matched between the 2 groups. For the propensity score–matched patient cohort, the detailed clinicopathologic characteristics and DRS were assessed and compared between the 2 groups. The association of each DRS category with the rate of structural recurrence during follow-up was then investigated and compared between the 2 groups.

Inclusion and exclusion criteria of robotic thyroid surgery for PTC

Before 2014, the inclusion criteria of our robotic thyroid surgery with a gasless transaxillary approach were limited to the following: (1) PTC with a maximum diameter of ≤ 2 cm and (2) minimal invasion to the anterior thyroid capsule and strap muscle. Since 2014, the inclusion criteria have been expanded to include a primary tumor size up to 4 cm. The exclusion criteria remain the same: definite posterior capsular invasion particularly adjacent to the tracheoesophageal groove, lateral LN metastasis, and distant metastasis at presentation. All patients were informed if they were candidates for RS based on the previously described inclusion and exclusion criteria. Next, they were informed about the costs, brief surgical procedures, possible complications, and expected outcomes of both RS and OS. Patients were then allowed to choose their preferred surgical procedure.

DRS system

All enrolled participants were divided into 4 categories (excellent, indeterminate, biochemically incomplete, and structurally incomplete), based on the response to the initial therapy 2 years after the completion of initial treatment, by using data obtained during the first 2 years of follow-up. The detailed definitions of each category were based on the per-treatment modalities in the DRS system proposed by Tuttle et al,[12] and are provided in Table 1.[10,11,14] The level of thyroglobulin (Tg) at a similar thyroid stimulating hormone (TSH) level was classified as stable, declining, or increasing after a >20% change in 2 consecutive measurements. Positive anti-Tg antibody (Ab) was defined as a serum anti-Tg Ab level of > 60 IU/mL. The trend in anti-Tg Ab change was also defined as stable, declining, or increasing by comparing 3 consecutive values.

Preoperative diagnosis and staging workup

The preoperative diagnosis was determined using ultrasonography-guided fine-needle aspiration cytology (FNAC) or core needle biopsy. In all cases of preoperative diagnosis of PTC, preoperative staging workup with both high-resolution neck ultrasonography and computed tomography (CT) was performed to evaluate tumor characteristics and cervical LN status.

Surgical strategy for PTC

During the study period, TT was recommended for almost all patients with primary tumors >1 cm and thyroid lobectomy was usually recommended for patients with tumors ≤1 cm, based on the 2009 ATA guidelines.

At our institution, central compartment node dissection (CCND) is routinely performed for patients with thyroid cancer; at a minimum, unilateral prophylactic CCND is performed even in cases without suspicious LNs, found on preoperative imaging studies or during surgery. Bilateral CCND is performed on patients with suspicious LNs or LN enlargement in the contralateral central compartment or in those with bilateral cancer. Compartment-based nodal dissection is performed according to preference and usually includes level VI with or without level VII. All surgeries on patients included in the present study were performed by a single endocrine surgeon (J.H.Y.).

RAI remnant ablation protocol

Subsequent RAI remnant ablation was performed 4–6 weeks after the initial operation, according to the protocol established by the Endocrinology Division of Asan Medical Center.[23] Briefly, the ablation doses used in this report were determined to be 30, 80, and 150 mCi, according to clinicopathologic parameters (ETE, resection margin, and cervical LN metastasis) at the time of surgery. Iodine-131 remnant ablation was not given for any patient with an intrathyroidal solitary tumor ≤1.0 cm. An ablative dose of 30 mCi was administered to patients with a multifocal tumor or a tumor >1.0 cm, without ETE or cervical LN metastasis. An ablative dose of 80 mCi was administered to patients with microscopic ETE or ipsilateral central cervical LN metastasis. An ablative dose of 150 mCi was administered to patients with gross ETE, positive surgical resection margin, or bilateral central cervical LN metastasis.

At the time of remnant ablation, after thyroid hormone withdrawal or recombinant human TSH administration, the serum stimulated Tg (sTg) (ablation sTg) level (reference range, 1.0–23.3 ng/mL) was measured with the anti-Tg Ab level (reference, < 60 U/mL) when the TSH level (reference range, 0.4–5.0 mU/L) was >30 mU/L. A postablation whole-body scan was performed 5–7 days after the administration of iodine-131.

Postoperative follow-up protocol

All patients underwent follow-up examinations at the outpatient clinic. Patients who underwent thyroid lobectomy alone or TT without RAI remnant ablation were followed up at 1, 6, and 12 months, and annually thereafter, while receiving thyroid hormone treatment to decrease their serum TSH to subnormal levels without thyrotoxicosis. Thyroid function tests, including serum

Table 1
Definition of each response-to-therapy category of the dynamic risk stratification.[10,11,14]

Response to therapy category	Definitions		
	TT with RAI ablation	TT without RAI ablation	Thyroid lobectomy alone
Excellent	Nonstimulated Tg < 0.2 ng/mL or stimulated Tg < 1.0 ng/mL and undetectable anti-TgAb levels and negative imaging	Nonstimulated Tg < 0.2 ng/mL or stimulated Tg < 2.0 ng/mL and undetectable anti-TgAb levels and negative imaging	Stable, nonstimulated Tg < 30 ng/mL and undetectable anti-TgAb levels and negative imaging
Indeterminate	Nonspecific findings on imaging studies or faint uptake in thyroid bed on RAI scanning or nonstimulated Tg 0.2–1 ng/mL or stimulated Tg 1–10 ng/mL or stable or declining anti-TgAb levels in the absence of structural or functional disease	Nonspecific findings on imaging studies or faint uptake in thyroid bed on RAI scanning or nonstimulated Tg 0.2–5.0 ng/mL or stimulated Tg 2–10 ng/mL or stable or declining anti-TgAb levels in the absence of structural or functional disease	Nonspecific findings on imaging studies or stable or declining anti-TgAb levels in the absence of structural or functional disease
Biochemical incomplete	Nonstimulated Tg > 1 ng/mL or stimulated Tg > 10 ng/mL or increasing anti-TgAb levels and negative imaging	Nonstimulated Tg > 5 ng/mL or stimulated Tg > 10 ng/mL or Increasing Tg values with similar TSH levels or Increasing anti-TgAb levels and negative imaging	Nonstimulated Tg[‡] > 30 ng/mL or increasing Tg values with similar TSH[‖] levels or increasing anti-TgAb levels and negative imaging
Structural incomplete	Structural or functional evidence of disease regardless of Tg or anti-TgAb levels	Structural or functional evidence of disease regardless of Tg or anti-TgAb levels	Structural or functional evidence of disease regardless of Tg or anti-TgAb

TT, total thyroidectomy; *RAI*, radioactive iodine; *Tg*, thyroglobulin; *TgAb*, thyroglobulin antibody; *TSH*, thyroid stimulating hormone.

TSH, Tg, and anti-Tg Ab levels, were routinely performed. Neck ultrasonography was performed once per year.

In patients who underwent TT with RAI remnant ablation, diagnostic whole-body scan after thyroid hormone withdrawal or recombinant human TSH administration was performed 6–12 months after remnant ablation, with the simultaneous measurement of serum sTg (control sTg). Serum Tg/anti-Tg Ab measurements and neck ultrasonography were performed on all patients during the follow-up period. When the control sTg was ≥1 ng/mL and neck ultrasonography showed no evidence of disease (NED), ^{18}F-deoxyglucose positron emission tomography (FDG-PET) or chest CT imaging were considered to localize persistent or remnant disease. Any patients suspected to have locoregional recurrence underwent ultrasonography-guided FNAC. Distant metastasis was diagnosed using whole-body scan, chest CT, or ^{18}F-FDG-PET/CT, and confirmed using serial imaging or biopsy. Patients who underwent thyroid lobectomy alone or TT without RAI remnant ablation were checked at 1, 7, and 12 months postoperatively, and every 12 months thereafter, while receiving thyroid hormone treatment to decrease their serum TSH to subnormal levels without thyrotoxicosis. Thyroid function tests, including those for serum TSH, Tg, and anti-Tg Ab levels, were routinely conducted at every visit, and neck ultrasonography was routinely performed once per year. For patients showing continuously increasing Tg or anti-Tg Ab levels on those consecutive measurements, although there was no identifiable structural lesion on neck ultrasonography, CT scans of the chest were performed to exclude the possibility of distant metastases. If the chest CT did not show any specific finding, further chest CT was not performed.

Structural recurrence was defined as the appearance of cytologically or histopathologically proven malignant tissue, or the appearance of highly suspicious structural lesions on cross-sectional or functional imaging studies after a period of NED for a minimum of 1 year after the initial treatment. However, unlike extracervical structural lesions (distant metastasis), even highly suspicious locoregional structural lesions on cross-sectional or functional imaging studies were not considered a structural recurrence until the lesion was confirmed using FNAC or surgical biopsy. Lesions not confirmed cytologically or histopathologically were re-evaluated using serial ultrasonography and repeated FNAC or surgical biopsy. Structural persistent disease was defined as the appearance of structural lesions without a period of NED for 1 year after the initial treatment. Biochemical recurrence, which was defined as an elevated serum Tg level (suppressed Tg ≥ 0.2 ng/mL and sTg ≥ 1 ng/mL in patients treated with TT) without clinical evidence of structural disease, was not classified as a true recurrence.

Statistics

Categorical variables are presented as absolute numbers and percentages; whereas continuous variables are presented as means ± standard deviations or as medians and ranges. Comparisons of clinicopathologic and treatment characteristics between the groups were performed using the *t* test for continuous data and the Fisher exact test for categorical data. Propensity scores were estimated with RS (or OS) as the dependent variable by using multiple logistic regression analysis. Model discrimination was assessed with C statistics (0.804), and model calibration was assessed with Hosmer-Lemeshow statistics ($\chi^2 = 8.992$, degree of freedom = 8, $P = .343$). Propensity score matching was performed using greedy matching with a caliper width of 0.2 standard deviations of the logit of the propensity score. The absolute standardized differences were used to determine the balance after matching. All absolute standardized differences after matching were <0.1, and the C statistic was 0.541 after matching. All *P* values were 2-sided, and $P < .05$ was considered statistically significant. Data were analyzed using SAS (v 9.4; SAS Institute, Cary, NC, USA).

Results

Comparison of clinicopathologic characteristics and DRS between the RS and OS groups before propensity score matching

The baseline clinicopathologic characteristics of all enrolled participants are provided in Table 2. The RS group showed a predominance of younger female patients, smaller primary tumor size, lower multifocality and bilaterality, less extensive ETE, less extensive thyroid resection performed, a higher proportion of patients undergoing RAI remnant ablation, and lower TNM stage than the OS group. However, in the subgroup analysis, according to initial treatment modalities, multifocality and bilaterality did not significantly differ between the 2 groups. Furthermore, tumor classification showed a significant difference only in patients treated with TT and RAI remnant ablation, probably owing to a higher proportion of T4a classification in the OS group than in the RS group.

With respect to the DRS, an excellent response was achieved in 71% of RS and 79.8% of OS, an indeterminate response in 22.3% of RS and 13.6% of OS, a biochemical incomplete response in 1.6% of

Table 2
Comparison of clinicopathologic characteristics between robotic and conventional open thyroidectomy groups before propensity score-matching.

Characteristics	Overall ($n=2112$)			Total with RAI* ($n=1069$)			Total without RAI* ($n=210$)			Lobectomy alone ($n=833$)		
	Robot ($n=492$)	Open ($n=1620$)	*P*	Robot ($n=141$)	Open ($n=928$)	*P*	Robot ($n=24$)	Open ($n=186$)	*P*	Robot ($n=327$)	Open ($n=506$)	*P*
Age, year, mean ± standard deviation	40.0 ± 8.9	50.3 ± 11.1	< .001	41.4 ± 8.7	51.4 ± 11.2	< .001	41.4 ± 8.6	54.2 ± 10.7	< .001	39.2 ± 9.1	46.9 ± 10.28	< .001
Gender, *n* (%)												
Female	449 (91.3)	1268 (78.4)	< .001	134 (95.0)	762 (82.1)	< .001	19 (79.1)	157 (84.4)	.555	291 (88.9)	350 (69.2)	< .001
Male	43 (8.7)	352 (21.7)		7 (5.0)	166 (17.9)		5 (20.9)	29 (15.6)		36 (11.1)	156 (30.8)	
Primary tumor size (cm)	0.7± 0.5	1.0 ± 0.9	< .001	1.1 ± 0.6	1.3 ± 0.9	.026	0.7 ± 0.5	0.7 ± 0.7	.969	0.8 ± 0.6	1.0 ± 1.1	.002
Multifocality, *n* (%)	109 (22.2)	540 (33.3)	< .001	61 (43.3)	421 (45.3)	.651	5 (20.9)	29 (15.6)	.555	43 (13.1)	90 (17.8)	.081
Bilaterality, *n* (%)	50 (10.2)	296 (18.3)	< .001	48 (34.0)	280 (30.2)	.378	2 (8.3)	16 (8.6)	.965	0	0	.999
Lymph node metastasis, *n* (%)	187 (38.0)	581 (35.8)	.265	66 (46.8)	423 (45.6)	.659	2 (8.3)	13 (7.0)	.766	119 (36.4)	145 (28.7)	.063
Extrathyroidal extension, *n* (%)												
Non-extensive	490 (99.6)	1551 95.7)	< .001	139 (98.6)	863 (92.9)	< .001	24 (100)	184 (98.9)	.087	327 (100)	504 (99.6)	<.001
Extensive	2 (0.4)	69 (4.3)		2 (1.4)	65 (7.1)		0	2 (1.1)		0	2 (0.4)	
Extent of thyroidectomy, *n* (%)			< .001									
Total thyroidectomy	165 (33.5)	1114 (68.8)										
Thyroid lobectomy	327 (66.5)	506 (31.2)										
RAI* ablation, *n* (%)	141 (85.5)	928 (83.3)	< .001	141 (85.5)	928 (83.3)	< .001						
Dosage of RAI[8] ablation, millicurie, *n* (%)												
30	60 (42.6)	403 (43.4)	.707	60 (42.6)	403 (43.4)	.707						
80	54 (38.3)	373 (40.2)		54 (38.3)	373 (40.2)							
150	27 (19.1)	152 (16.4)		27 (19.1)	152 (16.4)							
Tumor classification												
T1a	204 (41.5)	583 (36.0)	< .001	20 (14.2)	166 (17.9)	.007	21 (87.5)	156 (83.9)	.974	163 (49.8)	261 (51.5)	.239
T1b	33 (6.7)	116 (7.2)		20 (14.2)	89 (9.6)		1 (4.2)	8 (4.3)		12 (3.7)	19 (3.8)	
T2	3 (0.6)	38 (2.4)		0	23 (2.5)		0	1 (0.5)		3 (0.9)	14 (2.8)	
T3	250 (50.8)	814 (50.2)		99 (70.2)	585 (63.1)		2 (8.3)	19 (10.2)		149 (45.5)	210 (41.5)	
T4a	2 (0.4)	69 (4.2)		2 (1.4)	65 (6.9)		0	2 (1.1)		0	2 (0.4)	
Node classification												
N0	304 (61.8)	1039 (64.2)	.265	75 (53.2)	505 (54.4)	.659	22 (91.7)	173 (93.0)	.766	208 (63.6)	361 (71.3)	.063
N1a	187 (38.0)	581 (35.8)		66 (46.8)	423 (45.6)		2 (8.3)	13 (7.0)		119 (36.4)	145 (28.7)	
TNM* stage												
I	379 (77.0)	839 (51.8)	< .001	95 (67.4)	333 (35.9)	<.001	24 (100)	168 (90.3)	.393	260 (79.5)	338 (66.8)	<.001
II	0	16 (1.0)		0	8 (0.9)		0	0		0	8 (1.6)	
III	111 (22.6)	696 (43.0)		44 (31.2)	522 (56.3)		0	16 (8.6)		67 (20.5)	158 (31.2)	
IVa	2 (0.4)	69 (4.2)		2 (1.4)	65 (6.9)		0	2 (1.1)		0	2 (0.4)	
Dynamic risk stratification, *n* (%)												
Unavailable	19 (3.9)	59 (3.6)	<.001	6 (4.3)	37 (4.0)	.780	0	9 (4.8)	.529	13 (4.0)	13 (2.6)	.075
Excellent	350 (71.1)	1293 (79.8)		114 (80.9)	773 (83.3)		20 (83.3)	154 (82.8)		216 (66.1)	366 (72.3)	
Indeterminate	110 (22.3)	220 (13.6)		19 (13.4)	96 (10.3)		4 (16.7)	20 (10.8)		87 (26.6)	104 (20.6)	
Biochemical incomplete	8 (1.6)	36 (2.3)		1 (0.7)	13 (1.4)		0	3 (1.6)		7 (2.1)	20 (3.9)	
Structural incomplete	5 (1.1)	12.0 (0.7)		1 (0.7)	9 (1.0)		0	0		4 (1.2)	3 (0.6)	

* The 7th American Joint Committee on Cancer/Union for International Cancer Control Tumor Node Metastasis staging system.*RAI*, radioactive iodine.

RS and 2.3% of OS, and a structural incomplete response in 1.1% of RS and 0.7% of OS, showing significant differences between the 2 groups ($P < .001$; Table 2). However, in the subgroup analysis, according to initial treatment modalities, the DRS did not significantly differ between the 2 groups (Table 2).

Comparison of clinicopathologic characteristics and DRS between the RS and OS groups after propensity score matching

After propensity score matching with regard to age, sex, primary tumor size, ETE, multifocality, and bilaterality, 444 matched pairs of patients were compared in detail, as shown in Table 3. The median follow-up period was 60 months (range, 28–99 months). The extents of thyroidectomy ($P < .001$) and LN dissection ($P < .001$) were significantly lesser in the RS group than in the OS group. The RS group also showed a significantly higher proportion of patients undergoing RAI remnant ablation after TT than the OS group. However, the proportion of patients undergoing RAI remnant ablation with a high dose (150 mCi) was rather higher in the OS group than in the RS group. Although less extensive LN dissection was performed in the RS group, the incidence of LN metastasis, number of retrieved LNs, number of metastatic LNs, metastatic LN ratio, and maximal size of the metastatic foci of LNs did not differ significantly between the 2 groups. The TNM stage ($P=.297$) and ATA initial risk stratification ($P=.139$) were not statistically different between the 2 groups. The occurrence of structural

Table 3
Comparison of clinicopathologic characteristics between propensity score-matched robotic and conventional open thyroidectomy groups.

Characteristics	Overall ($N = 888$)			Total with RAI ($n = 342$)			Total without RAI ($n = 67$)			Lobectomy alone ($n = 479$)		
	Robot ($n = 444$)	Open ($n = 444$)	P	Robot ($n = 135$)	Open ($n = 207$)	P	Robot ($n = 23$)	Open ($n = 44$)	P	Robot ($n = 286$)	Open ($n = 193$)	P
Age, y, mean ± standard deviation	41.3 ± 8.3	41.3 ± 8.7	.999	42.2 ± 8.1	41.6 ± 9.1	.596	43.4 ± 7.0	42.9 ± 9.0	0.816	40.6 ± 8.4	40.4 ± 8.2	.792
Sex, *n* (%)												
Female	401 (90.3)	392 (88.2)	.385	128 (94.8)	191 (92.3)	.388	23 (100)	41 (93.2)	0.313	250 (87.4)	160 (82.9)	.168
Male	43 (9.7)	52 (11.8)		7 (5.2)	16 (7.7)		0	3 (6.8)		36 (12.6)	33 (17.1)	
Primary tumor size (cm)	0.8 ± 0.5	0.8 ± 0.4	.380	1.0 ± 0.6	0.9 ± 0.5	.144	0.6 ± 0.4	0.5 ± 0.3	0.351	0.7 ± 0.5	0.6 ± 0.3	.074
Extrathyroidal extension, *n* (%)												
Non-extensive	441 (99.3)	441 (99.3)	.999	133 (98.5)	204 (98.6)	.981	23	44	0.999	285 (99.7)	193	.999
Extensive	3 (0.7)	3 (0.7)		2 (1.5)	3 (1.4)		0	0		1 (0.3)	0	
Multifocality, *n* (%)	107 (24.1)	116 (26.1)	.536	61 (45.2)	76 (36.7)	.142	6 (26.1)	5 (11.4)	0.167	40 (14.0)	35 (18.1)	.249
Bilaterality, *n* (%)	50 (11.3)	43 (9.7)	.999	48 (35.6)	41 (19.8)	.002	2 (8.7)	2 (4.5)	0.603	—	—	—
LN metastasis, *n* (%)	162 (36.5)	163 (36.7)	.720	62 (45.9)	96 (46.4)	.616	3 (13.0)	3 (6.8)	0.548	97 (33.9)	64 (22.4)	.699
No. of retrieved LNs, *n*												
Ipsilateral CCND	6.4 ± 4.5	6.9 ± 5.2	.159	7.2 ± 4.6	7.9 ± 6.0	.345	5.0 ± 3.1	6.8 ± 5.1	0.194	6.6 ± 4.4	7.2 ± 5.2	.137
Bilateral CCND	8.0 ± 5.6	10.7 ± 8.3	.120	8.5 ± 4.6	11.2 ± 8.1	.054	6.5 ± 5.0	10.5 ± 5.7	0.343	3.5 ± 3.0	4.0 ± 3.5	.826
Number of metastatic LNs, *n*	1.1 ± 2.1	1.0 ± 2.0	.715	1.3 ± 2.0	1.3 ±1.9	.906	0.3 ± 0.8	0.1 ± 0.3	.142	0.8 ± 1.5	0.7 ± 1.3	.639
Metastatic LN ratio	0.15 ± 0.26	0.13 ± 0.22	.065	0.21 ± 0.31	0.17 ± 0.26	.186	0.05 ± 0.16	0.02 ± 0.11	.505	0.13 ± 0.24	0.10 ± 0.19	.263
Maximal size of metastatic LNs, cm	0.3 ± 0.2	0.3 ± 0.3	.974	0.4 ± 0.3	0.3 ±0.3	.436	0.6 ± 0.5	0.3 ± 0.5	.449	0.3 ± 0.2	0.3 ± 0.2	.567
Extent of thyroidectomy, *n* (%)												
Total thyroidectomy	158 (35.6)	251 (56.3)	<.001									
Lobectomy	286 (64.4)	193 (43.7)										
Extent of LN dissection, *n* (%)												
No	5 (1.1)	9 (2.0)	<.001	1 (0.7)	5 (2.5)	<.001	0	3 (6.8)	<.001	4 (1.4)	1 (0.5)	.293
Ipsilateral CCND	414 (93.2)	281 (63.3)		115 (85.2)	79 (38.1)		21 (91.3)	16 (36.4)		278 (97.2)	186 (96.4)	
Bilateral CCND	25 (5.7)	154 (34.7)		19 (14.1)	123 (59.4)		2 (8.7)	25 (56.8)		4 (1.4)	6 (3.1)	
RAI ablation, *n* (%)	135 (86.1)	207 (82.5)	.001	135 (86.1)	207 (82.5)	.001						
Dosage of RAI ablation, millicurie, *n* (%)												
30	66 (48.8)	90 (43.4)	<.001	66 (48.8)	90 (43.4)	<.001						
80	57 (42.5)	77 (37.5)		57 (42.5)	77 (37.5)							
150	12 (8.7)	40 (19.1)		12 (8.7)	40 (19.1)							
Tumor classification												
T1a	177 (39.9)	202 (45.4)	.328	19 (14.1)	58 (28.0)	.007	19 (82.6)	39 (88.7)	0.698	139 (48.6)	105 (54.4)	.398
T1b	31 (7.0)	30 (6.8)		18 (13.3)	24 (11.6)		1 (4.4)	2 (4.5)		12 (4.2)	4 (2.1)	
T2	3 (0.7)	6 (1.4)		0	5 (2.4)		0	0		3 (1.1)	1 (0.5)	
T3	231 (52.0)	203 (45.7)		96 (71.1)	117 (56.6)		3 (13.0)	3 (6.8)		132 (46.1)	83 (43.0)	
T4a	2 (0.4)	3 (0.7)		2 (1.5)	3 (1.4)		0	0		0	0	
Node classification												
N0	282 (63.5)	281 (63.3)	.720	73 (54.1)	111 (53.6)	0.616	20 (87.0)	41 (93.2)	.548	189 (66.1)	129 (77.6)	.699
N1a	162 (36.5)	163 (36.7)		62 (45.9)	96 (46.4)		3 (13.0)	3 (6.8)		97 (33.9)	64 (22.4)	
TNM* stage												
I	331 (74.6)	344 (77.5)	.297	89 (65.9)	139 (67.2)	0.960	23	44	.999	219 (76.6)	161 (83.4)	.074
II	0	2 (0.4)		0	1 (0.5)		0	0		0	1 (0.5)	
III	111 (25.0)	95 (21.4)		44 (32.6)	64 (30.9)		0	0		67 (23.4)	31 (16.1)	
IVa	2(0.4)	3 (0.7)		2 (1.5)	3 (1.4)		0	0		0	0	
ATA[] risk, *n* (%)												
Low	341 (76.8)	320 (72.1)	.139	85 (63.0)	134 (64.7)	.607	22 (95.6)	42 (95.5)	0.266	233 (81.5)	144 (74.6)	.176
Intermediate	90 (20.3)	101 (22.7)		40 (29.6)	53 (25.6)		0	2 (4.5)		51 (17.8)	46 (23.8)	
High	13 (2.9)	23 (5.2)		10 (7.4)	20 (9.7)		1 (4.4)	0		2 (0.7)	3 (1.6)	
Dynamic risk stratification, *n* (%)												
Excellent	375 (84.5)	355 (79.9)	.086	113 (83.8)	166 (80.2)	.487	19 (82.6)	37 (84.1)	0.999	243 (85.0)	152 (78.8)	.542
Indeterminate	57 (12.8)	81 (18.2)		20 (14.8)	40 (19.3)		4 (17.4)	7 (15.9)		33 (11.5)	34 (17.6)	
Biochemical incomplete	8 (1.8)	7 (1.6)		1 (0.7)	1 (0.5)		0	0		7 (2.4)	6 (3.1)	
Structural incomplete	4 (0.9)	1 (0.3)		1 (0.7)	0		0	0		3 (1.1)	1 (0.5)	
Follow-up duration, month	60 (28-95)	60 (28-99)	.693									
Structural recurrence, *n* (%)	7 (1.6)	2 (0.5)	.094	5 (3.7)	2 (1.0)	.118	0	0	.999	2 (0.7)	0	.518

* The 7th American Joint Committee on Cancer/Union for International Cancer Control Tumor Node Metastasis staging system;*LN*, lymph node; *CCND*, central compartment node dissection; *RAI*, radioactive iodine; *ATA*, American thyroid Association.

Table 4
Clinicopathological characteristics of patients with structural recurrence.

Sex	Patient 1 Female	Patient 2 Female	Patient 3 Male	Patient 4 Female	Patient 5 Female	Patient 6 Female	Patient 7 Female	Patient 8 Female	Patient 9 Female
Age at diagnosis	37	35	37	28	39	49	44	43	39
Surgery type	Open	Open	Robot	Robot	Robot	Robot	Robot	Robot	Robot
Thyroidectomy	Total	Total	Lobectomy	Lobectomy	Total	Total	Total	Total	Total
LN* dissection	Ipsilateral	Ipsilateral	Ipsilateral	Ipsilateral	Ipsilateral	Ipsilateral	Ipsilateral	Ipsilateral	Ipsilateral
Dosage of RAI† remnant ablation, millicurie	150	150	-	-	150	80	150	150	80
Location of structural recurrence	Ipsilateral Lateral LN	Ipsilateral Lateral LN	Contralateral Lobe	Contralateral Lobe	Ipsilateral Lateral LN	Ipsilateral Lateral LN	Contralateral central LN	Ipsilateral Lateral LN	Ipsilateral Lateral LN
Duration from initial treatment to structural recurrence, months	29	46	31	49	39	34	51	42	30
TNM* stage	I	I	I	I	I	III	I	I	I
ATA§ risk	Intermediate	High	Low	Intermediate	Low	Low	Low	Intermediate	High
DRS‖	Indeterminate	Indeterminate	Structural incomplete	Excellent	Excellent	Indeterminate	Excellent	Indeterminate	Structural incomplete
Anti-Tg Ab¶ status	Positive	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Positive
Control stimulated Tg†, ng/mL	–	1.50	–	–	0.08	0.08	0.53	1.90	–
Changing trend of Tg† at the time of DRS‖	–	Declining	Declining	Stable	Declining	Declining	Declining	Declining	–
Changing trend of anti-Tg Ab¶ at the time of DRS‖	Declining	–	–	–	–	–	–	–	Increasing

* TNM, the 7th American Joint Committee on Cancer/Union for International Cancer Control Tumor Node Metastasis staging system.
† Control stimulated Tg, serum thyroglobulin measured 6–12 months after remnant ablation at the time of a diagnostic whole-body scan after thyroid hormone withdrawal or Thyrogen administration only in patients with negative anti-thyroglobulin antibody.*LN*, lymph node; *RAI*, radioactive iodine; *ATA*, American thyroid association; *DRS*, dynamic risk stratification; *Anti-Tg Ab*, anti-thyroglobulin antibody.

Table 5
The reasons why the patients were classified into indeterminate or biochemical incomplete response to initial therapy after propensity score-matching.

Response-to-therapy category	Overall		Total with RAI		Total without RAI		Lobectomy alone	
	Robot	Open	Robot	Open	Robot	Open	Robot	Open
Indeterminate	57	81	20	40	4	7	33	34
Nonspecific findings on imaging studies, *n* (%)	32 (56.1)	22 (27.2)	11 (55.0)	9 (22.5)	1 (25.0)	2 (28.5)	20 (60.6)	11 (32.4)
Elevated nonstimulated Tg or stimulated Tg levels, *n* (%)	6 (10.5)	11 (13.6)	4 (20.0)	9 (22.5)	1 (25.0)	1 (14.3)	1 (3.0)	1 (2.9)
Stable or declining Anti-Tg Ab levels, *n* (%)	19 (33.4)	48 (59.2)	5 (25.0)	22 (55.0)	2 (50.0)	4 (57.2)	12 (36.4)	22 (64.7)
Biochemical incomplete	8	7	1	1	0	0	7	6
Elevated nonstimulated Tg or stimulated Tg levels, *n* (%)	1 (12.5)	2 (28.5)	1 (100)	1 (100)	0	0	0	1 (16.7)
Increasing Tg values with similar TSH levels, *n* (%)	5 (62.5)	3 (43.0)	0	0	0	0	5 (71.5)	3 (50.0)
Increasing anti-Tg Ab levels, *n* (%)	2 (25.0)	2 (28.5)	0	0	0	0	2 (28.5)	2 (33.3)

RAI, radioactive iodine; *Tg*, thyroglobulin; *anti-Tg Ab*, anti-thyroglobulin antibody; *TSH*, thyroid-stimulating hormone.

recurrence showed a higher tendency in the RS group than in the OS group; however, the difference between the 2 groups was not statistically significant ($P = .094$). The detailed clinicopathologic characteristics of patients with structural recurrence are provided in Table 4. Structural recurrences occurred in the contralateral lobe ($n = 2$), contralateral central compartment ($n = 1$), and lateral LN area ($n = 6$). Structural recurrence in the contralateral lobe ($n = 2$) occurred in patients treated with thyroid lobectomy alone, and the other structural recurrences occurred in those treated with TT and RAI remnant ablation.

With respect to the DRS, an excellent response was achieved in 84.5% of RS and 79.9% of OS, an indeterminate response in 12.8% of RS and 18.2% of OS, a biochemical incomplete response in 1.8% of RS and 1.6% of OS, and a structural incomplete response in 0.9% of RS and 0.3% of OS, showing no significant differences between the 2 groups ($P = .086$; Table 3). Even in the subgroup analysis according to initial treatment modalities, the DRS did not differ between the 2 groups (Table 3). Furthermore, the reasons why the patients were classified into an indeterminate or a biochemical incomplete response-to-therapy category in both the overall patients and the subgroups according to initial treatment modalities are provided in Table 5.

The incidence of structural recurrence based on each response-to-therapy category was also not statistically different between the 2 groups ($P = .278$; Table 6). The incidence of structural recurrence in patients with an excellent response to initial therapy was very low in both the RS and OS groups (Table 6).

Comparison of surgical completeness between propensity score–matched RS and OS groups

The surgical completeness between the 2 groups after propensity score matching was compared according to initial treatment modalities, with regard to the number of retrieved LNs, control sTg levels, and changing trend of anti-Tg Ab levels in patients who underwent TT and RAI remnant ablation, and with regard to the number of retrieved LNs, suppressed Tg levels, and changing trend of Tg or anti-Tg Ab levels in patients who underwent TT or thyroid lobectomy alone, showing no significant differences between the 2 groups (Table 7).

Discussion

This propensity score–matched comparison study between RS and OS focused on evaluating the DRS based on the response to initial therapy, and it was confirmed that the DRS was not significantly different between the 2 groups. This result can shed light on

Table 6
Incidence of structural recurrence based on each response-to-therapy category in the dynamic risk stratification.

		Structural recurrence		*P*
		Robotic, recurrence/overall (%)	Open, recurrence/overall (%)	
DRS	Excellent	3/375 (0.8)	0/355 (0)	.278
	Indeterminate	2/57 (3.5)	2/81 (2.5)	
	Biochemical incomplete	0/8 (0)	0/7 (0)	
	Structural incomplete	2/4 (50.0)	0/1 (0)	

DRS, dynamic risk stratification.

Table 7
Comparison of surgical completeness between propensity score–matched robotic and conventional open thyroidectomy groups.

Characteristics	Total with RAI ($n=342$)			Total without RAI ($n=67$)			Lobectomy alone ($n=479$)		
	Robot ($n=135$)	Open ($n=207$)	*P*	Robot ($n=23$)	Open ($n=44$)	*P*	Robot ($n=286$)	Open ($n=193$)	*P*
No. of retrieved LNs, *n*, mean ± standard deviation									
Ipsilateral CCND	7.2 ± 4.6	7.9 ± 6.0	.345	5.0 ± 3.1	6.8 ± 5.1	.194	6.6 ± 4.4	7.2 ± 5.2	.137
Bilateral CCND	8.5 ± 4.6	11.2 ± 8.1	.054	6.5 ± 5.0	10.5 ± 5.7	.343	3.5 ± 3.0	4.0 ± 3.5	.826
RAI[‡] ablation, *n* (%)	135 (86.1)	207 (82.5)	.001						
Anti-Tg Ab status, *n* (%)									
Negative	128 (94.8)	185 (89.4)	.077	21 (91.3)	40 (90.9)	.957	244 (85.3)	152 (78.8)	.106
Positive	7 (5.2)	22 (10.6)		2 (18.7)	4 (9.1)		42 (14.7)	41 (21.2)	
Control stimulated Tg[*], ng/mL, mean ± standard deviation	0.38 ± 1.48	0.19 ± 0.36	.084						
Control stimulated Tg[*] <1 ng/mL, *n* (%)	126 (93.3)	198 (95.7)	.056						
Suppressed Tg[†], ng/mL, mean ± standard deviation				0.21 ± 0.45	0.46 ± 1.57	.451	1.95 ± 2.84	2.01 ± 3.32	.825
Changing trend of suppressed Tg[†], *n* (%)						.138			.210
Declining				8 (34.8)	8 (18.2)		200 (69.9)	121 (62.7)	
Stable				15 (65.2)	32 (72.7)		81 (28.3)	66 (34.2)	
Increasing				0	4 (9.1)		5 (1.8)	6 (3.1)	
Changing trend of anti-Tg Ab[‡], *n* (%)			.431			.999			.069
Declining	6 (85.7)	21 (95.5)		2 (100)	4 (100)		31 (73.8)	35 (85.4)	
Stable	0	0		0	0		9 (21.4)	2 (4.8)	
Increasing	1 (14.3)	1 (4.5)		0	0		2 (4.8)	4 (9.8)	

[*] Control stimulated Tg, serum thyroglobulin measured 6–12 months after remnant ablation at the time of a diagnostic whole-body scan after thyroid hormone withdrawal or Thyrogen administration only in patients with negative anti-thyroglobulin antibody.
[†] Suppressed Tg, serum thyroglobulin levels and those changing trends under thyroid-stimulating hormone suppression therapy measured only in patients with negative anti-thyroglobulin level at the time of determining dynamic risk stratification.
[‡] Changing trend of anti-Tg Ab was evaluated only in patients with positive anti-thyroglobulin level at the time of determining dynamic risk stratification.*LN*, lymph node; *CCND*, central compartment node dissection; *RAI*, radioactive iodine; *anti-Tg Ab*, anti-thyroglobulin antibody.

the long-term oncologic outcome of RS for PTC; that is, it is likely to be comparable to that of OS.

Many previously published meta-analysis studies validated the postsurgical outcomes of RS for differentiated thyroid carcinoma, reporting reliable outcomes of RS regarding technical safety, quality of life, and surgical completeness, although those results somewhat varied according to each study population.[1–6] The technical safety and quality-of-life outcome of RS could be relatively well verified because the data required for evaluating those outcomes was cross-sectional and could be obtained during the early postoperative follow-up period. However, the long-term oncologic outcome of RS mainly for PTC is not yet well established because the median follow-up period of RS has not reached 10 years, which remains not even long enough to directly and accurately evaluate the oncologic outcome of RS with respect to structural recurrence.[7–9,24] Therefore, the oncologic safety of RS was evaluated and compared with that of OS with respect to surgical completeness by using postoperative serum sTg levels and the proportion of patients with sTg < 2 or 10 ng/mL at the time of RAI remnant ablation or sTg < 1 ng/mL at 6–12 months after the first ablation.[9,19–21,24–29] However, these measures can only be used for patients who underwent TT with RAI remnant ablation and had negative anti-Tg Ab levels, and therefore have limitations for the evaluation of patients treated without RAI remnant ablation or those who have positive anti-Tg Ab levels.

The DRS was designed to reclassify initial risk estimates determined by using histopathologic findings after the initial operation, based on the response to the initial therapy from data gathered from the first 2 years of follow-up. Therefore, the assessment of response-to-therapy categories in the DRS does not require long-term follow-up data. In addition, the DRS can be applied even in patients treated without RAI remnant ablation or those who have positive anti-Tg Ab, with some modified definitions of each response-to-therapy category.[11,14] Although the usefulness of DRS in patients treated without RAI remnant ablation is not as well evaluated as in those treated with RAI remnant ablation, previous studies validated DRS as a useful tool with a higher capability in predicting long-term structural recurrence than the ATA initial risk estimate or the American Joint Committee on Cancer/Union for Cancer Control TNM staging system.[12,13,30] In the present study, the DRS values based on response-to-therapy were significantly different between the RS and OS groups before propensity score matching, showing a lower proportion of excellent and biochemical incomplete responses and a higher proportion of indeterminate and structural incomplete responses in the RS group. However, grouping the “excellent response” and “indeterminate response” categories together as a “good response,” and “biochemical incomplete” and “structural incomplete” responses together as an “incomplete response,” the response-to-therapy classification in the RS group was similar to that in the OS group. Furthermore, in the subgroup analysis according to initial treatment modalities, the DRS did not significantly differ between the 2 groups. After propensity score matching, the DRS in patients receiving RS was comparable to that in patients who underwent OS, in both overall patients and subgroups according to initial treatment modalities. Of interest, an excellent response showing very low incidence

of structural recurrence was observed more frequently in the RS group than in the OS group, although the extent of thyroid surgery and LN dissection was less in the RS group than in the OS group. This might be associated with the biologic behavior of each cancer, which was not well established. DRS has been shown to reflect the biologic behavior of cancer and the response to initial therapy. Accordingly, although less extensive thyroid surgery or LN dissection was performed in the RS group for patients with the same clinicopathologic features, their responses to the initial therapy might be better than those of patients treated with more extensive OS according to the biologic behavior of cancers. In the present study, the RS group showed a higher tendency of developing structural recurrence than the OS group. However, all of the structural recurrences occurred in an undissected area at the time of the initial operation, and none of the patients with structural recurrence showed any evidence of lateral LN metastasis or bilaterality on preoperative staging workup with high-resolution neck ultrasonography and CT scan. Furthermore, various parameters that were evaluated to compare surgical completeness according to initial treatment modalities did not significantly differ between the RS and OS groups. Therefore, it makes sense to consider that the structural recurrences in the RS group did not result from incomplete resection and were unpreventable. Nevertheless, the higher tendency of structural recurrence in the RS group in the present study should be re-evaluated based on long-term follow-up results. To the best of our knowledge, the present study is the first to compare the short-term oncologic outcome of RS by using the DRS in a more comprehensive cohort of patients with PTC, with respect to the extent of thyroid surgery.

The present study has limitations inherent to its retrospective study design and exclusion criteria. Excluding patients with unavailable data for assessing response-to-therapy categories might be an important source of bias. In the present study, one of the main reasons for exclusion was missing Tg or anti-Tg Ab levels. At the time of data collection and analysis, we found many patients to have been checked only for free T4 and TSH, although this was not our follow-up strategy. The other reason was delayed measurement of Tg and anti-Tg Ab levels beyond 2 years after the initial operation, which made it impossible to evaluate the changing trends of Tg and anti-Tg Ab levels within 2 years after the initial operation. However, basic clinicopathologic data other than the DRS was not significantly different between patients with and without available data for assessing response-to-therapy categories (data not presented). We also tried to avoid selection bias through propensity score matching between the 2 groups. Theoretically, the appropriateness of propensity score matching was determined using the C statistic; the closer the C statistic to 1, the more appropriate the propensity score matching. On the basis of the C statistic of 0.804 in the present study, the propensity score was estimated to be well matched between the 2 groups. Our institutional preference for prophylactic CCND could also be a possible source of bias. However, owing to the preference for prophylactic LN dissection for PTCs at our institution, all of the LN-related factors could be evaluated in detail. Nevertheless, the indolent nature of PTC itself, the large proportion of patients with PTC $\leq$1 cm (for whom the 2015 ATA guidelines now recommend active surveillance as an alternative to immediate surgery), the small number of enrolled patients, and the short follow-up duration might mask the difference in outcomes between the RS and OS groups in the present study. Therefore, a further study based on a larger patient cohort and a longer follow-up duration will be needed to accurately compare the long-term oncologic outcomes between the 2 groups.

The DRS should be considered a useful tool in predicting the risk of long-term structural recurrence by using early postoperative data from approximately 2 years after the initial operation in patients undergoing RS who had a relatively short follow-up period. In the present study, an association of DRS with long-term oncologic outcome could not be clearly evaluated because of the short follow-up period. However, on the basis of the high predictability of DRS for long-term structural recurrence or persistent disease that has been described in the literature,[10–15,16] the long-term oncologic safety of RS is expected to be comparable to that of OS in properly selected patients with PTC, like our inclusion criteria for RS in the present study. In addition, the oncologic outcomes of RS for more advanced thyroid cancer can be compared preliminarily with that of OS by using the DRS.

References

1. Jackson NR, Yao L, Tufano RP, Kandil EH. Safety of robotic thyroidectomy approaches: Meta-analysis and systematic review. *Head Neck*. 2014;36:137–143.
2. Kandil E, Hammad AY, Walvekar RR, Hu T, Masoodi H, Mohamed SE, et al. Robotic thyroidectomy versus nonrobotic approaches: A meta-analysis examining surgical outcomes. *Surg Innov*. 2016;23:317–325.
3. Lang BH, Wong CK, Tsang JS, Wong KP, Wan KY. A systematic review and meta-analysis comparing surgically-related complications between robotic-assisted thyroidectomy and conventional open thyroidectomy. *Ann Surg Oncol*. 2014;21:850–861.
4. Pan JH, Zhou H, Zhao XX, Ding H, Wei L, Qin L, et al. Robotic thyroidectomy versus conventional open thyroidectomy for thyroid cancer: A systematic review and meta-analysis. *Surg Endosc*. 2017;31:3985–4001.
5. Son SK, Kim JH, Bae JS, Lee SH. Surgical safety and oncologic effectiveness in robotic versus conventional open thyroidectomy in thyroid cancer: A systematic review and meta-analysis. *Ann Surg Oncol*. 2015;22:3022–3032.
6. Wang YC, Liu K, Xiong JJ, Zhu JQ. Robotic thyroidectomy versus conventional open thyroidectomy for differentiated thyroid cancer: Meta-analysis. *J Laryngol Otol*. 2015;129:558–567.
7. Tae K, Ji YB, Cho SH, Lee SH, Kim DS, Kim TW. Early surgical outcomes of robotic thyroidectomy by a gasless unilateral axillo-breast or axillary approach for papillary thyroid carcinoma: 2 years' experience. *Head Neck*. 2012;34:617–625.
8. Tae K, Song CM, Ji YB, Sung ES, Jeong JH, Kim DS. Oncologic outcomes of robotic thyroidectomy: 5-year experience with propensity score matching. *Surg Endosc*. 2016;30:4785–4792.
9. Yi O, Yoon JH, Lee YM, Sung TY, Chung KW, Kim TY, et al. Technical and oncologic safety of robotic thyroid surgery. *Ann Surg Oncol*. 2013;20:1927–1933.
10. Momesso DP, Tuttle RM. Update on differentiated thyroid cancer staging. *Endocrinol Metab Clin North Am*. 2014;43:401–421.
11. Tarasova VD, Tuttle RM. A risk-adapted approach to follow-up in differentiated thyroid cancer. *Rambam Maimonides Med J*. 2016;7. doi:10.5041/RMMJ.10231.
12. Tuttle RM, Tala H, Shah J, Leboeuf R, Ghossein R, Gonen M, et al. Estimating risk of recurrence in differentiated thyroid cancer after total thyroidectomy and radioactive iodine remnant ablation: Using response to therapy variables to modify the initial risk estimates predicted by the new American Thyroid Association staging system. *Thyroid*. 2010;20:1341–1349.
13. Jeon MJ, Kim WG, Park WR, Han JM, Kim TY, Song DE, et al. Modified dynamic risk stratification for predicting recurrence using the response to initial therapy in patients with differentiated thyroid carcinoma. *Eur J Endocrinol*. 2014;170:23–30.
14. Momesso DP, Vaisman F, Yang SP, Bulzico DA, Corbo R, Vaisman M, et al. Dynamic risk stratification in patients with differentiated thyroid cancer treated without radioactive iodine. *J Clin Endocrinol Metab*. 2016;101:2692–2700.
15. Park S, Kim WG, Song E, Oh HS, Kim M, Kwon H, et al. Dynamic risk stratification for predicting recurrence in patients with differentiated thyroid cancer treated without radioactive iodine remnant ablation therapy. *Thyroid*. 2017;27:524–530.
16. Castagna MG, Maino F, Cipri C, Belardini V, Theodoropoulou A, Cevenini G, et al. Delayed risk stratification, to include the response to initial treatment (surgery and radioiodine ablation), has better outcome predictivity in differentiated thyroid cancer patients. *Eur J Endocrinol*. 2011;165:441–446.
17. Voutilainen PE, Siironen P, Franssila KO, Sivula A, Haapiainen RK, Haglund CH. AMES, MACIS and TNM prognostic classifications in papillary thyroid carcinoma. *Anticancer Res*. 2003;23:4283–4288.
18. Lang BH, Lo CY, Chan WF, Lam KY, Wan KY. Staging systems for papillary thyroid carcinoma: A review and comparison. *Ann Surg*. 2007;245:366–378.
19. Schlumberger M, Catargi B, Borget I, Deandreis D, Zerdoud S, Bridji B, et al. Strategies of radioiodine ablation in patients with low-risk thyroid cancer. *N Engl J Med*. 2012;366:1663–1673.
20. Giovanella L, Ceriani L, Suriano S, Ghelfo A, Maffioli M. Thyroglobulin measurement before rhTSH-aided 131I ablation in detecting metastases from differentiated thyroid carcinoma. *Clin Endocrinol*. 2008;69:659–663.
21. Nascimento C, Borget I, Troalen F, Al Ghuzlan A, Deandreis D, Hartl D, et al. Ultrasensitive serum thyroglobulin measurement is useful for the follow-up of patients treated with total thyroidectomy without radioactive iodine ablation. *Eur J Endocrinol*. 2013;169:689–693.
22. Edge SB, Compton CC. The American Joint Committee on Cancer: The 7th edition of the AJCC cancer staging manual and the future of TNM. *Ann Surg Oncol*. 2010;17:1471–1474.

23. Kim EY, Kim TY, Kim WG, Yim JH, Han JM, Ryu JS, et al. Effects of different doses of radioactive iodine for remnant ablation on successful ablation and on long-term recurrences in patients with differentiated thyroid carcinoma. *Nucl Med Commun*. 2011;32:954–959.
24. Lee SG, Lee J, Kim MJ, Choi JB, Kim TH, Ban EJ, et al. Long-term oncologic outcome of robotic versus open total thyroidectomy in PTC: A case-matched retrospective study. *Surg Endosc*. 2016;30:3474–3479.
25. Durante C, Montesano T, Attard M, Torlontano M, Monzani F, Costante G, et al. Long-term surveillance of papillary thyroid cancer patients who do not undergo postoperative radioiodine remnant ablation: Is there a role for serum thyroglobulin measurement? *J Clin Endocrinol Metab*. 2012;97:2748–2753.
26. Kim TY, Kim WB, Kim ES, Ryu JS, Yeo JS, Kim SC, et al. Serum thyroglobulin levels at the time of 131I remnant ablation just after thyroidectomy are useful for early prediction of clinical recurrence in low-risk patients with differentiated thyroid carcinoma. *J Clin Endocrinol Metab*. 2005;90:1440–1445.
27. Kim WG, Yoon JH, Kim WB, Kim TY, Kim EY, Kim JM, et al. Change of serum antithyroglobulin antibody levels is useful for prediction of clinical recurrence in thyroglobulin-negative patients with differentiated thyroid carcinoma. *J Clin Endocrinol Metab*. 2008;93:4683–4689.
28. Mallick U, Harmer C, Yap B, Wadsley J, Clarke S, Moss L, et al. Ablation with low-dose radioiodine and thyrotropin alfa in thyroid cancer. *N Engl J Med*. 2012;366:1674–1685.
29. Webb RC, Howard RS, Stojadinovic A, Gaitonde DY, Wallace MK, Ahmed J, et al. The utility of serum thyroglobulin measurement at the time of remnant ablation for predicting disease-free status in patients with differentiated thyroid cancer: A meta-analysis involving 3947 patients. *J Clin Endocrinol Metab*. 2012;97:2754–2763.
30. Pitoia F, Bueno F, Urciuoli C, Abelleira E, Cross G, Tuttle RM. Outcomes of patients with differentiated thyroid cancer risk-stratified according to the American thyroid association and Latin American thyroid society risk of recurrence classification systems. *Thyroid*. 2013;23:1401–1407.