Peer Review File

Article information: https://dx.doi.org/10.21037/gs-21-105

Reviewer A: This is a retrospective study demonstrating the risk levels re-stratified due to postoperative findings, which suggests a multidisciplinary approach in preoperative evaluation. Although most of the findings provided in this article is agreeable, some of the aspects of this study needs additional consideration.

Comment 1: According to the authors, all patients were included with low risk WDTC in a 13 years period (Line 17-18). However, 301 patients are questionably low concerning the high incidence of thyroid cancer, which dramatically elevated in the last few decades. **Additional description regarding the patient inclusion process should be presented, specifically in the number of patients excluded and the reason why**.

Response 1: We thank the reviewer for bringing this into our attention. Following this comment, we added description regarding the patient inclusion process and regarding the patients excluded (see page 5 lines 2-14). Also, we added a flow chart of inclusion process (see **Figure 1**):

"Patients were defined as being at low risk for disease-specific recurrence if they had well-differentiated thyroid tumors between 1-4 cm in size, free of any evidence of positive cervical lymph nodes, invasion to adjacent structures, and high-risk cytology. Clinically positive nodes were defined as being abnormal by US findings. We excluded patients not eligible for initial hemi-thyroidectomy according to the National Comprehensive Cancer Network (NCCN) and the ATA thyroid cancer management guidelines. Thus, Patients with thyroid tumors that were not well-differentiated, or with well-differentiated thyroid tumors that were smaller than 1 cm or larger than 4 cm in size, were excluded from the study. Other exclusion criteria included patients with other malignancies except for the thyroid malignancy, and patients with preoperatively known high risk characteristics, according to preoperatively imaging and clinical examination, such as gross extra-thyroidal extension on preoperative imaging, clinically apparent cervical lymph node metastases, distant metastases, vocal cord paralysis or immobility on physical examination, history of radiation and positive family history (**Figure 1**)".

Comment 2: According to Table 1, there were several patients presenting with symptoms such as hoarseness, dyspnea, and dysphagia, which are very unusual concerning low risk WDTC cases. **Please comment the portion of the patients that were associated with thyroid pathology, and if not, further explanation of the cause of these symptoms should help the readers for further understanding**.

Response 2: We thank the reviewer for this comment. We agree with the reviewer statement that symptoms such as hoarseness, dyspnea and dysphagia are unusual in patients with low risk WDTC. However, the mentioned patients did report those symptoms. It is important to mention that our cohort included patients who were admitted at our department for surgery and were asked specifically regarding the initial complaints that have led them to medical consultation. Thus, any subjective complaint they reported was documented, even if there was no objective explanation for the cause

of these symptoms. Also, important to mention, that those reported symptoms were mild, with no supporting findings in the physical examination.
On the basis of our collected data, we cannot estimate the portion of patients with complaints of hoarseness, dyspnea or dysphagia who underwent thyroid US, eventually had thyroid pathology that is perhaps incidental. This is an interesting question however not in the scope of our study.
Following the reviewer comment, we added clarification to the manuscript (see page 6, lines 16-23):
"Of note, the presenting symptoms reported by some patients (such as hoarseness, dyspnea and dysphagia) are unusual in patients with low risk WDTC. However, our cohort included patients who were admitted at our department for surgery and were asked specifically regarding the initial complaints that have led them to medical consultation. Thus, any subjective complaint they reported was documented, even if there was no objective explanation for the cause of these symptoms. Also, important to mention, that those reported symptoms were mild, with no supporting findings in the physical examination".

Comment 3: According to the pre-operative risk stratification of eighth AJCC edition, the age factor is critical in determining the risk of WDTC (R Michael Tuttle, Ali S Alzahrani, Risk Stratification in Differentiated Thyroid Cancer: From Detection to Final Follow-Up, The Journal of Clinical Endocrinology & Metabolism, 2019;104, (9):4087–4100.). Younger patients under age 55, are classified as low risk (stage I) regardless of tumor size, lymph node status or ETE. Thus, even if ETE was present postoperatively, as long as distant metastasis is absent, it could still be classified as low-risk cancer (even though risk assessment for recurrence might point to a higher risk). However, it is hard to catch this kind of incidence from the manuscript. **It would be clearer if it is explained in detail concerning the age group in table 2, further distinguishing the upscaled factors by two age groups (under and over 55 years old) separately**.
Response 3: We thank the reviewer for this comment.
We are aware of the fact that according to the pre-operative risk stratification of eighth AJCC edition, the age factor is critical in determining the risk of WDTC, as the reviewer mentioned (R Michael Tuttle, Ali S Alzahrani, Risk Stratification in Differentiated Thyroid Cancer: From Detection to Final Follow-Up, The Journal of Clinical Endocrinology & Metabolism, 2019;104, (9):4087–4100). Younger patients under age 55, are classified as low risk (stage I) regardless of tumor size, lymph node status or ETE. Thus, even if ETE was present postoperatively, as long as distant metastasis is absent, it could still be classified as low-risk cancer (even though risk assessment for recurrence might point to a higher risk).
In our study, we referred to **low risk for disease-specific recurrence**, as mentioned throughout the manuscript (page 1 line 7, page 1 lines 15-17, page 4 lines 8-9, page 5 line 2).
Following this comment, we separated in **table 3** between the two age groups (under and over 55 years old).
We re-analyzed our data and examined whether there were upscaled factors

distinguishing the two age groups (under and over 55 years old).
According to our results, among patients younger than 55 years only microscopic positive margins were found in a significant rate in the upscaled group (n=20 (12.9% vs n=11 (44%), P=0.01). However, among the patients older than 55 years, not only microscopic positive margins were found to be more common among the upscaled patients but also older age and larger size of the lesion.
Following these findings, we added our conclusions to the manuscript (see page 8 lines 20-24 and page 9 lines 1-3):
"The eighth AJCC edition (www.cancerstaging.org) defined the age factor as critical in determining the pre-operative risk stratification of WDTC[7,8].
Accordingly, we further analyzed our data according to different age groups (**Table 3**). 121(40.2%) patients were older than 55 years of age. The microscopic positive margins were found to be significantly more common in the upscaled specimens of both younger and older groups (20 (12.9%) vs 11 (44%), P=0.001); 17 (17%) vs 9 (42.9%), P=0.007; accordingly). Interestingly in the older age group both patient age (63.9±6.9 vs 68.6±7.9, P=0.032) and tumor size (1.5±0.64 vs 1.9±0.7) were found to be predictors of tumor upscaling".
Accordingly, we added the following paragraph to the discussion section (see page 10 lines 7-16):
"In October 2016, the AJCC (www.cancerstaging.org) published the eighth edition of the AJCC/TNM cancer staging system, revising thyroid cancer staging. Among the modifications made, was an increase of the age cutoff from 45 years to 55 years of age at diagnosis and including age as a major factor determining prognostic stage groups and the risk of WDTC preoperatively[7,8]. Accordingly, we performed a subgroup analysis to the two age groups (younger and older than 55 years of age) comparing low risk patients who were upscaled to higher risk group, to those who were not upscaled. Interestingly, only in the older age group patients' age and tumor size were predictors of upscaling. Thus, we recommend considering a more aggressive approach towards this subgroup, discussing the possible risks of partial thyroidectomy in the presence of a large tumor in an older patient (**Table 3**)".
Table 3. Difference in demographic, clinical, intra operative and final histopathology among patients who were upscaled vs those who did not, sub-grouped according to the eighth AJCC age cut-off (older and younger than 55 years)

Reviewer B
General comments: This is a retrospective review of patients operated on for WDTC between 2006-2018. They were included based on the retrospectively assigned ATA low-risk grading and the proportion of patients who would theoretically have then been upstaged based on intra-operative findings or pathological examination was calculated.
There are a number of issues that need to be addressed and in its current form, the manuscript is not deemed suitable for acceptance.
Please see the following general and specific comments:

The central node data needs to be reported, especially given that it is a major point of

difference to other studies that have reported on post-operative upstaging.

Comment 1: How many patients had a central lymph node dissection prophylactically (if any)
Response 1: None of the patients in our study had a prophylactic central lymph node dissection. We added this information to our manuscript (see page 5 line 15 and also page 14 line 15-16): "None of the patients included had undergone a prophylactic central lymph node dissection".

Comment 2: How many patients had nodal tissue in their specimens? **How many of these were involved?**
Response 2: Six of our patients (2%) had accidental nodal tissue detected in their final histopathology specimens. Of these, **only in one patient** (0.3%), a microscopic single lymph was detected as involved. This finding didn't alter the patient's risk stratification. Following the reviewer comment, we added this information to our manuscript in the results section (see page 8 lines 11-19):
"Regarding the central neck compartment, important to mention that none of the patients included had undergone a prophylactic central lymph node dissection. Since we have excluded patients who underwent neck dissection, all lymph nodes detected in the 301 patients' cohort, were found incidentally on histopathologic examination. Six of those 301 patients (2%) had incidental lymph nodes detected in their final histopathology specimens. None of the patients had more than 5 positive lymph nodes removed. Of these, only in one patient (0.3%), a single positive microscopic lymph was detected. This finding didn't alter the patient's risk stratification. Moreover, none of these six patients with incidental lymph nodes detected in their final histopathology specimens was upscaled regarding risk stratification, thus none of them was included in the 46 upscaled patients' cohort".

Comment 3: This then needs to be compared to the data of the comparative studies.
Response 3: We searched the literature for data regarding incidental nodal tissue detected in thyroid surgeries without central lymph node dissection.
We found that in the study of Kluijfhout et al., the incidence of incidental nodal tissue detected in thyroid surgeries was mentioned (reference number 20 in our manuscript: Kluijfhout WP, Pasternak JD, Lim J, Kwon JS, Vriens MR, Clark OH, Shen WT, Gosnell JE, Suh I, Duh QY. Frequency of High-Risk Characteristics Requiring Total Thyroidectomy for 1-4 cm Well-Differentiated Thyroid Cancer. Thyroid. 2016;26(6):820-4).
In their retrospective analysis, Kluijfhout et al. sought to determine how often a completion TT would be recommended based on the 2015 ATA guidelines if lobectomy was initially performed in patients with 1-4 cm WDTC without preoperatively known risk factors. According to their results, of 1000 consecutive patients operated for WDTC, 287 (29%) would have been eligible for lobectomy as the initial operation according to the recent NCCN and ATA guidelines. The mean age in this cohort was 45 years, and 80% were women. Aggressive tall cell variant histology was found in 1 patient (0.5%), angio-invasion in 34 (12%), ETE in 48 (17%), positive margins in 51 (18%), and **positive lymph**

nodes in 49 patients (17%). Completion TT would have been recommended in 122/287 (43%) patients. Even in those with 1- 2 cm cancers, completion TT would have been recommended in 52/143 (36%). They concluded that nearly half of the patients with 1-4 cm WDTC who are eligible for lobectomy under current guidelines would require completion TT based on pathological characteristics of the initial lobe.

At the end of their discussion, Kluijfhout et al. mention the limitations of their study. One of the limitations mentioned, is that **they used any positive lymph nodes within the specimen as high-risk characteristic, whereas patients with ≤ 5 lymph nodes (smaller than 0.2 cm) are still considered ATA low risk**. Since they didn't perform prophylactic neck dissections and all patients with clinically N1 were excluded, none of their patients had more than 5 lymph nodes removed. It is possible that of the patients without available size of the lymph node metastasis, some contained micrometastatic (smaller than 0.2 cm) disease and would have been considered ATA low risk. Thus, they state, that the rates of completion TT in their study may be overestimated.

In our study, none of the patients included had undergone a prophylactic central lymph node dissection. Since we have excluded patients who underwent neck dissection, all lymph nodes detected in the 301 patients' cohort, were found incidentally on histopathologic examination. Six of those 301 patients (2%) had incidental lymph nodes detected in their final histopathology specimens. None of the patients had more than 5 positive lymph nodes removed. Of these, only in one patient (0.3%), a single positive microscopic lymph was detected. This finding didn't alter the patient's risk stratification. Moreover, none of these six patients with incidental lymph nodes detected in their final histopathology specimens was upscaled regarding risk stratification, thus none of them was included in the 46 upscaled patients' cohort.

Accordingly, we modified our discussion section (see page 13, lines 16-23, page 14 lines 1-23):

"In their retrospective analysis, Kluijfhout et al.[20] sought to determine how often a completion thyroidectomy would be recommended based on the 2015 ATA guidelines if lobectomy was initially performed in patients with 1-4 cm WDTC without preoperatively known risk factors. They reviewed 1000 patients operated for WDTC and found that 287 (29%) would have been eligible for lobectomy as the initial operation according to the recent NCCN and ATA guidelines. In their study. nearly half of the patients with 1-4 cm WDTC who were eligible for lobectomy according to current NCCN and ATA guidelines (122/287, 43%) required completion thyroidectomy based on the final postoperative histopathological characteristics. In their study[20], incidental positive lymph nodes were found in 17% of patients eligible for HT (49/287). Of note, we report only 2 out of 46 (4%) patients that were upscaled due to the nodal status (patients with macroscopic lymph node involvement were of course not included). The discrepancy may be explained by the fact that our clinical preoperative evaluation is done meticulously by experienced ultrasonographists, cytologists, endocrinologists and head and neck surgeons together in a multidisciplinary dedicated team. Moreover, at the end of their discussion, Kluijfhout et al.[20] mention the limitations of their study. One of the limitations mentioned, is that they used any positive lymph nodes within the specimen as high-risk characteristic, whereas patients with ≤ 5 lymph nodes (smaller than 0.2 cm)

are still considered ATA low risk. Since they didn't perform prophylactic neck dissections and all patients with clinically N1 were excluded, none of their patients had more than 5 lymph nodes removed. It is possible that of the patients without available size of the lymph node metastasis, some contained micrometastatic (smaller than 0.2 cm) disease and would have been considered ATA low risk. Thus, they state, that the rates of completion TT in their study may be overestimated.
In our study, none of the patients included had undergone a prophylactic central lymph node dissection. Since we have excluded patients who underwent neck dissection, all lymph nodes detected in the 301 patients' cohort, were found incidentally on histopathologic examination. Six of those 301 patients (2%) had incidental lymph nodes detected in their final histopathology specimens. None of the patients had more than 5 positive lymph nodes removed. Of these, only in one patient (0.3%), a single positive microscopic lymph was detected. This finding didn't alter the patient's risk stratification. Moreover, none of these six patients with incidental lymph nodes detected in their final histopathology specimens was upscaled regarding risk stratification, thus none of them was included in the 46 upscaled patients' cohort".

Comment 4: **Please include complication** data – specifically rates of RLN injury.
Response 4: In accordance with the reviewer's comment, we added complication data to the results section (see page 9 lines 7-11):

"The surgery's' complications in our cohort were as follows: 1 patient (0.3%) had dehiscence of the surgical scar that resolved spontaneously by secondary healing. Ten patient (3.2%) had transient hypocalcemia. Injury to the recurrent laryngeal nerve occurred in 7 paitents (4.6%) - five patients (1.7%) had transient unilateral vocal fold impairment, and two patient (0.66%) had permanent unilateral voal fold paralysis. ".

Comment 5: Rajjoub et al.'s re-examination of the SEER data has not been discussed and is relevant to aspects of the discussion
o Rajjoub SR, Yan H, Calcatera NA et al. Thyroid lobectomy is not sufficient for T2 papillary thyroid cancers. Surgery 2018; 163: 1134–43
Response 5: In accordance with the reviewer's recommendation, we discussed this study in the discussion (see page 12, lines 7-19):
"Another factor that affects prognosis and should be taken into consideration when deciding on the extent of surgery is the histologic subtype of papillary thyroid cancer. In their study, Rajjoub et al.[18] aimed to examine whether survival is affected by extent of surgery for conventional vs. follicular-variant papillary thyroid cancer when stratified by tumor size. They evaluated 33,816 adults undergoing surgery for papillary thyroid cancer from 2004 to 2008 for 1.0-3.9 cm tumors and clinically negative lymph nodes. A total of 30,981 patients had TT and 2,835 had HT; 22,899 patients had conventional papillary thyroid cancer and 10,918 had follicular-variant papillary thyroid cancer. TT was associated with improved survival for conventional (P = 0.02) but not for follicular-variant papillary thyroid cancer patients (P = 0.42). For conventional papillary thyroid cancer, adjusted analysis showed TT was associated with improved survival for 2.0-3.9 cm tumors (P = 0.03) but not for 1.0-1.9 cm tumors (P = 0.16). For follicular-variant, HT

and TT had equivalent survival for 1.0-1.9 cm (P = 0.45) and 2.0-3.9 cm (P = 0.88) tumors".

Comment 6: Similarly, the following systematic review should be included in the discussion

o Chan S, Karamali K, Kolodziejczyk A et al. Systematic review of recurrence rate after Hemithyroidectomy for low-risk well-differentiated thyroid Cancer. Eur. Thyroid J. 2020; 9: 73–84

Response 6: In accordance with the reviewer's recommendation, we added this systematic review into the discussion (see page 11 lines 12-23, page 12 lines 1-3):

"In their systematic review of recurrence rate and survival after HT for low-risk WDTC, Chan et al[17]. identified 31 studies (with a total of 228,746 patients (HT: 36,129, TT: 192,617), which had published recurrence and/or survival data for patients having had HT for WDTC. Pooled recurrence rates were 9.0% for HT (which is significantly higher than in previously published reports) compared to 7.4% for TT, (odds ratio, OR, 1.45; CI 1.16–1.81, *p* =0.001). Further, this rate was maintained when examining patients within low-risk cohorts established with recognised risk classifications (AGES, MACIS, AMES, AJCC). Subgroup analysis demonstrated a pooled recurrence rate of 9.2% for HT and 5.3% for TT. They also discovered that of those patients who develop recurrent disease, 48% recur outside the central neck. Pooled 10-year overall survival rates were similar - 95.7% for HT and 95.8% for TT (OR 0.92; CI 0.73–1.18, *p* = 0.52), consistent with current opinion that overall survival in low-risk WDTC is favourable independent of surgical extent.

Although their findings indicate that there is a small but significantly higher recurrence rate after HT compared to TT, the evidence base was heterogenous and subject to confounding factors and would ultimately benefit from prospective randomized trials to overcome these deficiencies".

Comment 7: Can results be stratified into 2 cohorts – one pre and one post the introduction of the 2015 guidelines?

Response 7: We thank the reviewer for this comment.

Following this comment, we stratified the results into 2 cohorts - one pre and one post the introduction of the 2015 guidelines published in 2016 (see **Table 4**).

Accordingly, we added the following paragraph to the results section (page 9 lines 4-6):

"In order to clarify the effect of the ATA 2015 guidelines (published in 2016) on the amount and characteristics of upscaled patients, we further divided our cohort into patients upscaled prior to the publication of the 2015 ATA guidelines, and those upscaled after the publication (**Table 4**)".

Also, we added the following paragraph to the study limitations at the end of the discussion (page 16 lines 8-10 and page 16 lines 18-21):

"…the majority of our cohort was operated prior to 2015, therefore we only had limited data regarding the effect of the recent 2015 ATA guidelines on patients' risk assessment and upscaling assessment".

"Only 14 patients were operated after the 2015 ATA guidelines, of which only 2 were

upscaled. No significant predictors of upscaling were detected apart from tumor size. We noted that among patients operated prior to the application of the 2015 ATA guidelines, larger tumors were more common in the upscaled group (1.6 ±0.7 vs 1.9±0.8 cm, P=0.039)".
Table 4. the demographic, clinical intra operative and histopathological characteristics of low-risk patients who were operated before and after the introduction of 2015 ATA guidelines.

Comment 8: Please consider referring to "partial thyroidectomy" as "hemi-thyroidectomy" or "lobectomy". Hemi-thyroidectomy implies total hemi-thyroidectomy, whereas the term 'partial' may be considered to mean a subtotal resection of one of the lobes.
Response 8: We thank the reviewer for bringing this into our attention. Accordingly, we changed the term "partial thyroidectomy" to "hemi-thyroidectomy" throughout the manuscript.

Specific Comments:
Introduction:
Comment 9: Line 6 – please insert 'a' ("has a good overall prognosis")
Response 9: We inserted "a".

Comment 10: Line 12: "while preserving adequate minimal risk of recurrence" – revise wording
Response 10: We changed the wording, to: "without increasing risk of recurrence and metastasis".

Comment 11: Line 19: "… risk for" should be "Risk of"
Response 11: We changed to "risk of" as recommended.

Methods
Comment 12: Line 16 – "… ethical committee" should be "ethics committee"
Response 12: We changed to "ethics committee" as recommended.

Results
Comment 13: Please consider reporting the results with two significant figures – ie 15.28% as 15%
Response 13: Following the reviewer recommendation, we presented the results throughout the manuscript with only two significant figures.

Discussion
Comment 14: Line 9 "showed that determinate survival for all patients" – please rephrase
Response 14: Following the reviewer recommendation, we rephrased this sentence:
"…showed overall survival of 87% at 10 years".

Comment 15: How many of your patients had central nodes in the pathology specimen and how many had formal central node dissection performed as part of the operation? Please comment on this in comparison to the proportion reported by Kluijfhout et al. – these needs to be explored if reasons for the differences in nodal upstaging are to be speculated on.

Response 15: None of the patients in our study had a prophylactic central lymph node dissection. We added this information to our manuscript (see page 8 lines 11-12, page 14 lines 15-16): "None of the patients included had undergone a prophylactic central lymph node dissection".

Six of our patients (2%) had accidental nodal tissue detected in their final histopathology specimens. Of these, **only in one patient** (0.3%), a microscopic single lymph was detected as involved. This finding didn't alter the patient's risk stratification. Following the reviewer comment, we added this information to our manuscript in the results section (see page 8 lines 11-19):

"Regarding the central neck compartment, important to mention that none of the patients included had undergone a prophylactic central lymph node dissection. Since we have excluded patients who underwent neck dissection, all lymph nodes detected in the 301 patients' cohort, were found incidentally on histopathologic examination. Six of those 301 patients (2%) had incidental lymph nodes detected in their final histopathology specimens. None of the patients had more than 5 positive lymph nodes removed. Of these, only in one patient (0.3%), a single positive microscopic lymph was detected. This finding didn't alter the patient's risk stratification. Moreover, none of these six patients with incidental lymph nodes detected in their final histopathology specimens was upscaled regarding risk stratification, thus none of them was included in the 46 upscaled patients' cohort".

As recommended by the reviewer, we comment on this in comparison to the proportion reported by Kluijfhout et al. Accordingly, we modified our discussion section (see page 13, lines 16-23, page 14 lines 1-23):

"In their retrospective analysis, Kluijfhout et al.[20] sought to determine how often a completion thyroidectomy would be recommended based on the 2015 ATA guidelines if lobectomy was initially performed in patients with 1-4 cm WDTC without preoperatively known risk factors. They reviewed 1000 patients operated for WDTC and found that 287 (29%) would have been eligible for lobectomy as the initial operation according to the recent NCCN and ATA guidelines. In their study. nearly half of the patients with 1-4 cm WDTC who were eligible for lobectomy according to current NCCN and ATA guidelines (122/287, 43%) required completion thyroidectomy based on the final postoperative histopathological characteristics. In their study[20], incidental positive lymph nodes were found in 17% of patients eligible for HT (49/287). Of note, we report only 2 out of 46 (4%) patients that were upscaled due to the nodal status (patients with macroscopic lymph node involvement were of course not included). The discrepancy may be explained by the fact that our clinical preoperative evaluation is done meticulously by experienced ultrasonographists, cytologists, endocrinologists and head and neck surgeons together in a multidisciplinary dedicated team. Moreover, at the end of their

discussion, Kluijfhout et al.[20] mention the limitations of their study. One of the limitations mentioned, is that they used any positive lymph nodes within the specimen as high-risk characteristic, whereas patients with ≤ 5 lymph nodes (smaller than 0.2 cm) are still considered ATA low risk. Since they didn't perform prophylactic neck dissections and all patients with clinically N1 were excluded, none of their patients had more than 5 lymph nodes removed. It is possible that of the patients without available size of the lymph node metastasis, some contained micrometastatic (smaller than 0.2 cm) disease and would have been considered ATA low risk. Thus, they state, that the rates of completion TT in their study may be overestimated.
In our study, none of the patients included had undergone a prophylactic central lymph node dissection. Since we have excluded patients who underwent neck dissection, all lymph nodes detected in the 301 patients' cohort, were found incidentally on histopathologic examination. Six of those 301 patients (2%) had incidental lymph nodes detected in their final histopathology specimens. None of the patients had more than 5 positive lymph nodes removed. Of these, only in one patient (0.3%), a single positive microscopic lymph was detected. This finding didn't alter the patient's risk stratification. Moreover, none of these six patients with incidental lymph nodes detected in their final histopathology specimens was upscaled regarding risk stratification, thus none of them was included in the 46 upscaled patients' cohort".

Comment 16: Overall, the discussion lacks flow and would benefit from revision to enhance clarity.
Response 16: We thank the reviewer for this comment. Following this comment, we revised the discussion to enhance clarity (see discussion section).

Reviewer C: The authors present a well-designed, single institution retrospective review to try to determine how many patients with low-risk thyroid cancer undergoing surgery would be re-stratified to intermediate or high-risk disease after the initial resection according to ATA guidelines. Their study included 301 patients and spanned 12 years. They found that only about 15% of patients would be re-stratified to higher risk disease and thus might need a completion thyroidectomy if they initially received a thyroid lobectomy. The aim is stated clearly and appropriately addressed. I have only minor comments I hope will strengthen the paper.

Comment 1: The authors appropriately observe that since many of the included patients were treated before the new ATA guidelines, the rate of total thyroidectomy is likely higher in the study than would be seen with current practices. That being said, some patients might still receive a total thyroidectomy for other factors- contralateral nodules, concern for gross invasion intra-op, compressive symptoms etc. Therefore, even if 15% are upscaled in terms of risk, not all will require a completion thyroidectomy assuming some would have had an initial total thyroidectomy. **Can the authors estimate (which may prove too difficult retrospectively) how many might have received an initial total thyroidectomy or how many in their cohort could reasonably have been expected to need a completion thyroidectomy?**

Response 1: We thank the reviewer for the comment. We stress that patient selection was performed to include only patients that were eligible for lobectomy according to the 2015 ATA guidelines, and only intra-operative and post-operative characteristics were examined to assess the effect of the latter on post-operative re stratification. Therefore, theoretically all patients included were eligible for lobectomy and accordingly, after final histopathological assessment, 85% could have maintained their initial low risk stratification.
As mentioned in the discussion section (page 9, lines 18-21):
"In total, 255 (85%) patients retained their preoperative low-risk level and thus were eligible for lobectomy according to the ATA 2015 guidelines. Forty-six patients were upscaled postoperatively from low risk to an intermediate-to-high risk level yielding a rate of postoperative risk escalation of 15%".
Also, in the discussion section, we mentioned among the limitations of our study, the following limitation (see page 16 lines 8-17):
"Second, the majority of our cohort was operated prior to 2015, therefore we only had limited data regarding the effect of the recent 2015 ATA guidelines on patients' risk assessment and upscaling assessment. The majority of our patients were operated in light of the ATA 2009 guidelines[33] thus a higher rate of TT was performed as compared to the rate that could have been performed in light of the ATA 2015 guidelines. The latter may lead to an internal bias due to the fact that the whole specimen was available for pathological evaluation. Moreover, the fact that the majority of our cohort was operated according to the ATA 2009 guidelines limited our ability to define the clinical implications of our findings regarding the upscaling rate. Thus, the data regarding follow-up of our cohort's patients is out of the scope of this current study".

Comment 2: Page 9, line 21- please clarify "clinical Neck findings".
Response 2: "clinical neck findings" refers to the presence of US detected abnormal lateral neck lymph nodes (cN1). Accordingly, we rephrased this in the manuscript (see page 13 lines 6-8):
"…initially clinically N-positive (clinical N1 disease; cN1), with presence of US detected abnormal lateral neck lymph nodes".

Comment 3: Table 1- Please verify the p values for hoarseness, dyspnea, and vocal cord paralysis are correct. The p value for vocal cord paralysis should not be zero. The p values for hoarseness and dyspnea should be exact, not <1.
Response 3: We modified table 1 accordingly (see **Table 1**).

Comment 4: Abstract, line 19- "aslow" should be " as low"
Response 4: We thank the reviewer for bringing this in to our attention. We corrected accordingly.