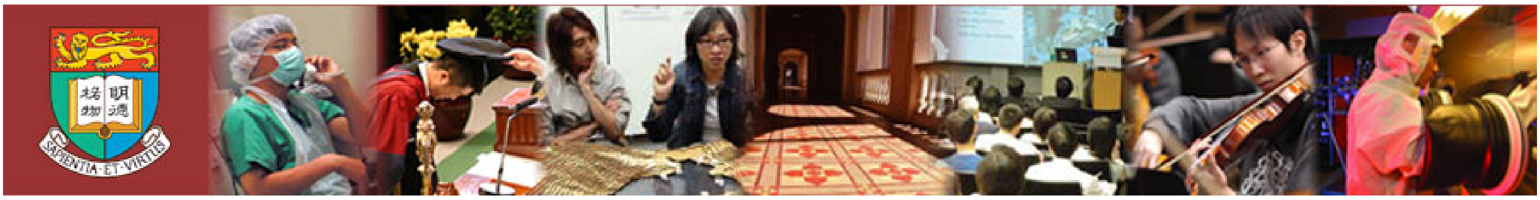

Title	A systematic review and meta-analysis evaluating completeness and outcomes of robotic thyroidectomy
Author(s)	Lang, HHB; Wong, CKH; Tsang, JS; Wong, KP; Wan, KY
Citation	The Laryngoscope, 2015, v. 125 n. 2, p. 509-518
Issued Date	2015
URL	http://hdl.handle.net/10722/205973
Rights	The Laryngoscope. Copyright © John Wiley & Sons, Inc.

SYSTEMATIC REVIEW

A systematic review and meta-analysis evaluating completeness and outcomes of robotic thyroidectomy

Brian Hung-Hin LANG[1], MS, FRACS

Carlos KH WONG[2], PhD

Julian Shun TSANG[1], MBBS, MRCS

Kai Pun WONG[1], MBBS, MRCS

Koon Yat WAN[4], MBBS, FRCR

[1]Department of Surgery, The University of Hong Kong, Hong Kong SAR, China

[2]Department of Family Medicine and Primary Care, University of Hong Kong, 3/F Ap Lei Chau Clinic, 161 Main Street, Ap Lei Chau, Hong Kong

[3]Department of Clinical Oncology, The University of Hong Kong, Hong Kong SAR, China

Address for Correspondence:

Dr Brian HH Lang

Division of Endocrine Surgery, Department of Surgery, Queen Mary Hospital, 102 Pokfulam Road, Hong Kong SAR, China

Tel.: (852) 22554232; Fax No.: (852) 28172291; Email: blang@hku.hk

Keywords: robotic thyroidectomy; papillary thyroid carcinoma; radioactive iodine, thyroglobulin; central lymph nodes; total thyroidectomy

Running title: Total thyroidectomy by robot is less complete.

Level of evidence: 3a

FINANCIAL DISCLOSURE: no financial disclosures

CONFLICT OF INTERESTS: none

ABSTRACT

Background: Despite immense interest, robotic-assisted thyroidectomy (RT) remains controversial in differentiated thyroid carcinoma (DTC). This systematic review and meta-analysis compared surgical completeness and/or oncological outcomes between RT and open thyroidectomy (OT) in low-risk DTC.

Methods: A systematic review was performed to identify studies that compared surgical completeness and/or oncological outcomes between RT and OT in DTCs. Any study that compared at least one parameter relating to surgical completeness and /or oncological outcome for DTC was considered. Number of central lymph nodes (CLNs) retrieved during central neck dissection (CND), pre-ablation stimulated thyroglobulin (sTg) level, radioiodine uptake on post-therapy scan and locoregional recurrence (LRR) were examined. Meta-analysis was performed using a fixed or random-effect model depending on heterogeneity between studies.

Results: Ten studies were eligible. Of the 2205 DTCs, 752 (34.1%) had RT while 1453 (65.9%) had OT. Relative to OT, RT had significantly fewer CLNs retrieved during CND (4.7 ± 3.2 vs. 5.5 ± 3.8, SMD=-0.240,95%CI=-0.364 - -0.116, p<0.001) and higher pre-ablation sTg level (3.6 ± 6.7ng/mL vs. 2.0 ± 5.0ng/mL, SMD=0.272,95%CI=0.022 – 0.522,p=0.033). Interestingly, these differences were more evident in the robotic transaxillary approach (RTAA) than the robotic bilateral axillo-breast approach. After a mean follow-up of 17.7 months, no LRR was found in RT while after 18.6 months, 1 LRR was found in OT.

Conclusions: Relative to OT, total thyroidectomy by RTAA was associated with fewer CLNs retrieved and less-complete thyroid resection. However, using RTAA is unlikely to compromise the outcomes of low-risk DTC because of its inherently good prognosis.

INTRODUCTION

Thyroidectomy is a common surgical procedure and the standard cervical open thyroidectomy (OT) has been proven safe and effective under experienced hands [1]. However, to further improve the cosmetic result and increase patient satisfaction, endoscopic approaches with incisions made outside the neck (also known as the extracervical approaches) were developed [2,3]. In 2007, a South Korean group pioneered using the *da Vinci* robot (the so-called "robotic-assisted thyroidectomy" or RT) to improve the ergonomics of these extracervical approaches [4]. Currently, the two most-described approaches are the robotic transaxillary approach (RTAA) and the robotic bilateral axillo-breast approach (RBA) [3]. Despite the higher cost, the proponents view that RT offers improved motion of endoscopic instruments, clearer stereoscopic visual field and dampening of physiologic tremors [4]. Since the first report in 2009,[5] there has been immense interest both in the US and other parts of the world with various groups publishing their initial successful experience [6-9]. However, RT remains controversial. In October 2011, the Food and Drug Administration in the US revoked the approval on the use of the robot in thyroidectomy and parathyroidectomy [10].

Although recent meta-analyses have found that both RT and OT may have comparable surgical medium to long term outcomes,[11-13] it remains unclear whether RT could provide a similar level of surgical completeness and oncological outcome as the gold standard OT in the treatment of differentiated thyroid carcinoma (DTC). This is relevant because in some parts of the world, this procedure is done mostly for low-risk DTC [4,5]. To our knowledge, there have been a few single-institution studies specifically evaluating these aspects [14,15]. However, as with any meta-analysis, one major advantage of pooling data from various studies is that at times it could highlight interesting and important findings which may not be apparent with individual study due to the lack of power. Therefore, we conducted a systematic review and meta-analysis to compare surgical completeness and/or oncological outcomes between RT and OT.

METHODS

This systematic review and meta-analysis was conducted in accordance with the PRISMA statement[16].

Search strategy

Studies comparing surgical completeness +/- oncological outcomes between patients with DTC who underwent RT and OT were retrieved from the Scopus, Medline (PubMed) and Cochrane Library electronic databases on 20th February 2014. We used the following free text search terms in “All fields”

#1: ‘robotic thyroidectomy’

#2: ‘robotic assisted thyroidectomy’

#3: ‘robot thyroidectomy’

#4: #1 OR #2 OR #3

There was no language restriction or methodological filters. The bibliographies of two previous meta-analyses on RT were searched for other additional relevant references [11,12].

Study selection

All titles identified by the search strategy were screened independently by three authors (BHL, JST, KPW). Search results were compared, and disagreements were resolved by consensus. Abstracts of potentially relevant titles were then reviewed for eligibility and full-length articles were selected for closer examination. Any prospective or retrospective study that compared at least one variable on surgical completeness and /or oncological outcome for DTC was considered. These variables included number of central lymph nodes (CLNs) retrieved during surgery, postoperative stimulated thyroglobulin (sTg) levels, findings on post-therapy radioiodine scan (RxWBS) and locoregional recurrence (LRR). However, we excluded case reports, editorials, expert opinions, reviews without original data, studies on pediatric population, studies comparing outcomes between RT and endoscopic (i.e. non-robotic) thyroidectomy and studies evaluating patients undergoing robotic-assisted lateral neck dissection. For studies with limited data on sTg, LRR and follow-up duration, the corresponding author of those studies was individually contacted for further information. Multiple reports of the same dataset were assessed and the most representative or updated report of a study was included.

Data extraction

All data were extracted onto a standardized form. The primary data extracted from each article included: type or design of study, first authorship, country of origin, year of publication, patient demographics, patient selection for RT and OT, extent of surgery (total thyroidectomy (TT) or less than total thyroidectomy (LTT)), tumor characteristics such as histological type, presence of extrathyroidal extension, multicentricity and bilaterality, tumor size, CLN metastasis, number of CLNs retrieved, TNM stage, pre-ablative sTg, percentage of radioactive iodine uptake (RAIU) on RxWBS, locoregional recurrence (LRR) and follow-up duration (months). TT included near-TT, TT and TT with central neck dissection (CND) whereas LTT included hemithyroidectomy and subtotal thyroidectomy.

Study quality

The quality of all non-randomized studies was assessed using the Newcastle–Ottawa Scale (NOS)[17]. The NOS considers methods of patient selection, comparability of the study groups and reporting of important outcomes. A maximum rating of 9 may be given to individual studies. Studies achieving a rating of 5 or more were considered higher quality. Two reviewers (BHL, JST) independently assessed the quality of the studies and disagreement was resolved by consensus.

Statistical methods

For comparison of dichotomous variables between RT and OT, chi-square tests and Fisher's exact tests were used. Student t-test was used for comparison of continuous variables. All the individual outcomes were integrated with the meta-analysis software Review Manager Software 5.0 (Cochrane Collaborative, Oxford, England). Standardized mean differences (SMD) were calculated for tumor size, number of CLNs retrieved, postoperative sTg level and RAIU. Odds ratios (OR) were examined for other categorical characteristics and outcomes. Heterogeneity was assessed with Cochran's Q statistic (chi-square), with a p value <0.10 for significance, and with the I^2 test.[18] An I^2 value >50% was considered substantial heterogeneity. Results were aggregated and analyzed using a fixed-effect model if I^2 value was not ≤50%. For I^2 values >50%, results were analyzed using a random-effect model. Publication bias was estimated by Begg's rank correlation test and Egger's regression test [19]. The meta-analyses in this study were conducted using IBM SPSS Version 20.0 for Window and Comprehensive Meta-Analysis Version 2.2.064 (Biostat, Inc.).

RESULTS

Figure 1 shows the flowchart of studies retrieved and excluded. Of the 485 titles initially identified from the database search, 20 full-length articles were assessed for inclusion, of which 10 were excluded and 10 studies [14,15,20-27] were determined to be eligible and were included in this systematic review. Table 1 lists these 10 excluded articles [28-37] and the reason for their exclusion. No additional study was found from our search of the two bibliographies in previous meta-analyses.[11,12] Of these 10 articles excluded, 5 were excluded mainly because their data were superseded by later studies.[19,22-24]

Patient selection

Ultrasonography was used as routine preoperative imaging modality in all studies.[14,15,20-27] The inclusion and exclusion criteria for RT or OT were similar. Inclusions included patient age between 21 – 65 years old, size of DTC ≤ 2 – 4 cm, thyroid lobe size ≤ 6cm and body mass index ≤ 36.[14,15,20-27] Exclusions included previous neck irradiation, presence of lateral lymph node and distant metastases, and posteriorly located carcinoma.[20,23,24,26] In terms of selecting for RT or OT, 4 studies were based on patient preference[14,20,25,27] while the other 6 studies did not specify their selection method.[15,21-24,26] In one study,[21] patients in the OT group were selected in reverse, chronological consecutive order from the time when the robot was first implemented (i.e. historical controls) while the others were cohort in design.

Baseline characteristics

Table 2a shows a comparison of the baseline characteristics between the 10 eligible studies. There were no randomized trials. Eight studies were retrospective while two were prospective. Of the 2205 patients included, 752 (34.1%) had RT (RT group) while 1453 (65.9%) had OT (OT group). In the RT group, there were 522 TTs and 230 LTTs while in the OT group, there were

1235 TTs and 218 LTTs. Overall, the OT group had significantly a higher TT:LTT ratio than the RT group (1235:218 vs. 522:230, p<0.001). Eight studies evaluated RTAA [14,20-25,27] while 2 studies evaluated RBA[15,26]. Of the 10 studies, two were from the United States while the rest came from South Korea.

Age at operation was matched in 3 studies [20,21,24] while sex ratio was matched in 4 studies [20,21,24,25]. Overall, patients in RT were significantly younger (40.8 ± 8.6 vs. 50.3 ± 10.7 years old, p<0.001) and more likely females (644:50 vs. 1192:186, p<0.001) than OT. In terms of histological types, 8 studies[14,15,20,22,23,25-27] exclusively had PTC while 2 had both PTC and FTC [21,24]. Overall, FTC only accounted for 0.2% of the entire DTC cohort.

After checking the heterogeneity using p-value of $\chi 2$ (p-value of <0.1) and I^2, five characteristics/outcomes were considered to have significant heterogeneity between studies and they were tumor multicentricity, $\chi 2$=19.167, df=6, p=0.004, I^2=68.696%; bilaterality, $\chi 2$=13.652, df=6, p=0.034, I^2=56.050%; *TNM* stage, $\chi 2$=19.077, df=6, p=0.004, I^2=68.549%; sTg, $\chi 2$=21.954, df=5, p=0.001, I^2=77.225%; Post Scan, $\chi 2$=140.373, df=2, p<0.001, I^2=98.575%.

Tumor characteristics

Table 2b shows a comparison of tumor characteristics between RT and OT. There were no significant differences in the proportion of extrathyroidal extension (348/582 (59.8%) vs. 710/1203 (59.0%),OR=0.963, 95%CI=0.774 – 1.197, p=0.732), tumor multi-centricity (200/668 (29.9%) vs. 438/1291 (33.9%), OR=0.822, 95%CI=0.535 – 1.261, p=0.368) and bilaterality (121/442 (27.4%) vs. 276/1076 (25.7%), OR=1.092, 95%CI=0.675 – 1.766, p=0.721) between the two groups. The mean primary tumor size in RT was significantly smaller than OT (7.7 ± 3.3mm vs. 8.6 ± 4.6mm) (SMD=-0.198; 95%CI=-0.297 - -0.098, p<0.001). Among the 7 studies reporting the number of CLNs retrieved during surgery, unilateral CND was performed in all 7 studies [14,20,22,23,25-27]. Figure 2a shows the forest plot for number of CLNs retrieved. The mean number of CLNs retrieved in RT was significantly less than OT (4.7 ± 3.2 vs. 5.5 ± 3.8) (SMD=-0.240, 95%CI=-0.364 - -0.116, *p*<0.001). Interestingly, this difference was more noticeable in the RTAA subgroup analysis (4.7 ± 3.3 vs. 5.7 ± 4.0, SMD=-0.283, 95%CI=-0.412 - -0.154, *p*<0.001) than the RBA group (4.7 ± 2.7 vs. 4.8 ± 2.8, SMD=-0.036, 95%CI=-0.325 – 0.253, *p*=0.806), although only one study was included in the RBA group. Potential publication bias was not significant, as confirmed by the Begg analysis (Kendall's tau = 0.267, *p*=0.452) and the Egger regression test (z= 0.590, *p*=0.587). However, the incidence of CLN metastasis was comparable between RT and OT (37.1% vs. 36.9%, OR=1.035, 95%CI=0.842 – 1.272, *p*=0.745). In terms of *TNM* stages, the ratio for Stages I&II / Stages III&IV tumors for RT and OT were 2.49 and 1.03, respectively (OR=0.456, 95%CI=0.285 – 0.729, *p*=0.001) and proportion of stage I tumors was significantly more in RT than OT (71.2% vs. 50.2%, *p*<0.001).

Oncological outcomes

Table 3 shows a comparison of RAI ablation, postoperative sTg levels, post-therapy findings and LRR between RT and OT. Percentages of RAI ablation were given in 5 studies [14,15,23,24,27]. The percentage of RAI ablation in RT ranged from 25.0% to 100% while RAI ablation in OT ranged from 36.7% to 100%. Their dose ranged from 2.78 to 5.55 GBq. Overall, the proportion receiving RAI ablation was comparable between RT and OT (329/360 (92.7%) vs. 821/945 (87.7%) (OR=0.952, 95%CI=0.574 – 1.579, *p*=0.849). Three studies [15,24,27] specifically provided and compared the percentage of RAIU at the RxWBS. Two studies [15,24] found comparable RAIU at RxWBS while one recent study found RAIU at RxWBS was significantly higher in the RT group (0.10 vs. 0.05, *p*=0.002) [27].

Six studies [14,15,22,23,26,27] compared pre-ablation sTg levels between the two groups. One study [23] also had post-ablation sTg levels while another study [24] only reported non-stimulated Tg levels. Figure 2b shows the forest plot for pre-ablation sTg. The overall mean pre-ablation sTg was significantly higher in RT than that in OT (3.6 ± 6.7ng/mL vs. 2.0 ± 5.0ng/mL, SMD=0.272, 95%CI=0.022 – 0.522, p=0.033). When only RTAA was considered, the difference became more significant between the two groups (6.1 ± 9.0ng/mL vs. 2.5 ± 5.8ng/mL, SMD=0.428, 95%CI=0.276 – 0.580, p<0.001) while when only RBA was considered, sTg appeared comparable (1.2 ± 3.4ng/mL vs. 1.1 ± 2.7ng/mL, SMD=0.040, 95%CI=-0.122 – 0.202, p=0.632). After a mean follow-up of 17.7 ± 8.7 months in RT group, no recurrence was found in RT group and after a mean follow-up of 18.6 ± 8.8 months, 1 recurrence was found in the OT group. There was no significant difference in mean follow-up between the two groups (p=0.860).

Sensitivity analysis

Since there were significant differences in study design and size, sensitivity analyses were performed. Findings on extrathyroidal extension, multicentricity, bilaterality, tumor size, CLN metastases, TNM stages, RAI and RxWBS were similar when the two prospective studies [26,27] were excluded or when the two small studies [21,24] were excluded from the pooled data.

DISCUSSION

Although RT may have the benefits of superior cosmesis and improved patient satisfaction,[2-4] it remains a controversial procedure in the West [10] Apart from the higher procedural cost, surgical completeness and oncological safety with RT have not been fully addressed [10,32] Although previous studies have demonstrated comparable completeness between RT and OT, they were single-institution studies.[14,15] Our data showed that RT seemed to be less effective in complete removal of thyroid tissue. This point should be kept in mind when a surgeon has to decide for the type of operation especially in high-risk groups such as older males, or patients with a posteriorly-located carcinoma or history of radiotherapy.

Despite the increasing number of publications comparing outcomes between the two procedures, all eligible studies were non-randomized and were subjected to selection biases. This was evident when patient baseline characteristics were being pooled. Patients in RT were significantly younger (40.8 years old vs. 50.3 years old, $p<0.001$) and more likely to be female (92.8% vs. 86.5%, $p<0.001$). Also overall the tumors in RT were significantly smaller (7.7 ± 3.3mm and 8.6 ± 4.6mm, $p<0.001$) and earlier staged ($p<0.001$). Therefore, based on these findings, tumors in RT belonged to a better risk group and so to some extent, RT was expected to have a more complete surgical resection and better oncological outcome than OT. However, our data shows that RT may be associated with a less complete resection than OT. Of the 6 studies [14,15,22,23,26,27] comparing pre-ablation sTg levels, RT had significantly higher mean level than OT (3.6ng/mL vs. 2.0ng/mL, $p<0.001$) implying that greater amount of residual thyroid tissue was left after surgery. Interestingly, this difference was only found in the RTAA group (6.1ng/mL vs. 2.5ng/mL, $p<0.001$) and not in the RBA group (1.2ng/mL vs. 1.1ng/mL, $p=0.632$).

One study also found that the percentage of RAIU on RxWBS after RTAA was also significantly higher than that of the OT (0.10 vs. 0.05, p=0.002) [27]. These findings are in discordant to previous single-institution studies which found the surgical completeness was comparable between RT and OT [14,15] However, it is worth noting that this was only for patients who underwent TT because both sTg and RAIU only reflect the amount of residual thyroid remnant after TT and not LTT. We postulate that one reason why TT by RTAA might be associated with a less complete thyroid resection is because during RTAA, the surgeon often has difficulty completely removing the contralateral lobe (or the non-tumorous side) when the incision is placed in the opposite axilla [7]. Therefore, we think the most likely source of pre-ablation sTg after RTAA is from the contralateral thyroid remnant and not from the ipsilateral lobe or residual tumor tissue. However, in terms of oncological clearance, we do not believe RTAA would compromise the oncological outcome. This is especially when most of these procedures were done for papillary microcarcinoma and low-risk DTC and so leaving small amount of non-tumorous thyroid tissue is unlikely to compromise prognosis when the oncological outcome is already excellent. This is supported by the fact that in one particular study, despite the higher initial pre-ablation level, the post-ablation sTg levels became comparable after 9-12 months (1.73ng/mL vs. 1.55ng/mL, p=0.661) [23].

Another measurement of surgical completeness was the number of CLNs retrieved during concomitant CND. RT had significantly fewer CLNs retrieved during CND than OT (4.7 vs. 5.5, p<0.001). Although the absolute difference was small (<1 CLN), it was nevertheless statistically significant. Interestingly, this difference was again more evident in RTAA (4.7 vs. 5.7, p<0.001) than RBA (4.7 vs. 4.8, p=0.806). We believe this might be attributed to a combination of factors. One factor is the difference in surgical approach because RTAA is essentially a lateral approach while OT is a medial or midline approach. Nevertheless, since the incidence of CLN metastasis was comparable (37.1% vs. 36.9%, *p*=0.745), we do not believe that RTAA had significantly under-staged the nodal status of DTCs at the time of operation. However, it was difficult to know if these concomitant CNDs were done prophylactically or therapeutically as none of the studies stated this. Regarding the comparison of actual oncological outcome, given that there was only one LRR detected in both groups in a relatively short follow-up period, it was difficult to exactly know if there is really a difference in outcome between RT and OT.

Despite some interesting findings, there are several shortcomings which need acknowledgement. Firstly, since all eligible studies were non-randomized or of low quality (≤4 by NOS), selection biases could have accounted for some of the differences in surgical completeness. Furthermore, it is unclear whether some studies were truly measuring the same outcome (e.g. sTg or RxWBS). Secondly, the number of studies eligible for inclusion was relatively small. Since there was significant overlap in dataset between studies and one-fourth of eligible studies were excluded (see Table 1). Nevertheless, excluding the two smallest studies did not affect our results. Thirdly, given the very good prognosis (and low risk for LRR) in DTCs, a much larger cohort with significant longer follow-up is necessary to fully evaluate the oncological outcome of RT. Nevertheless, if one considers pre-ablation sTg level and the percentage of RAIU on RxWBS as surrogates for surgical completeness, it would appear that TT via the RTAA has less complete thyroid resection than TT via the open approach. Lastly, although assessment of publication bias was performed, non-significant p-values do not necessarily imply no publication bias as the number of included studies was relatively small (i.e. low power to detect a real difference).

Conclusion

Despite the significantly fewer number of CLNs retrieved in CND and less complete TT by RTAA, RTAA is unlikely to compromise the oncological outcome of patients with low-risk DTC. However, given the excellent prognosis with few recurrences in low-risk DTC, a much larger patient cohort with long prospective follow up is necessary to fully evaluate the oncological outcome of RT in the future.

ACKNOWLEDGMENTS

None

AUTHORS CONTRIBUTIONS

BHH Lang / CKH Wong / JS Tsang / KP Wong / KY Wan were involved in the review of literature, acquisition of data and drafting and completing the manuscript. BHH Lang / CKH Wong / JS Tsang were also involved in the review of literature and drafting the manuscript. BHH Lang / CKH Wong / JS Tsang / KP Wong conceived the study, participated in the co-ordination and the acquisition of data and helped to draft the manuscript. All authors read and approved the final manuscript.

REFERENECES

1. The Surgical Outcomes Monitoring & Improvement Program (SOMIP) report volume 3 (July 2010 – June 2011). Coordinated by the Quality and Safety Division of the Hong Kong Hospital Authority. Accessed on 15th October 2013. Available: http://www.ha.org.hk/visitor/ha_index.asp.

2. Lang BH. Minimally invasive thyroid and parathyroid operations: surgical techniques and pearls. Adv Surg. 2010;44:185-98.

3. Lang BH, Lo CY. Technological innovations in surgical approach for thyroid cancer. J Oncol. 2010;2010. doi:pii: 490719. 10.1155/2010/490719.

4. Chung WY. Pros of robotic transaxillary thyroid surgery: its impact on cancer control and surgical quality. Thyroid. 2012;22(10):986-7.

5. Kang SW, Jeong JJ, Nam KH, Chang HS, Chung WY, Park CS. Robot-assisted endoscopic thyroidectomy for thyroid malignancies using a gasless transaxillary approach. J Am Coll Surg. 2009;209(2):e1-7.

6. Lewis CM, Chung WY, Holsinger FC. Feasibility and surgical approach of transaxillary robotic thyroidectomy without CO(2) insufflation. Head Neck. 2010;32(1):121-6

7. Lang BH, Chow MP. A comparison of surgical outcomes between endoscopic and robotically assisted thyroidectomy: the authors' initial experience. Surg Endosc. 2011;25(5):1617-23

8. Berber E, Heiden K, Akyildiz H, Milas M, Mitchell J, Siperstein A. Robotic transaxillary thyroidectomy: report of 2 cases and description of the technique. Surg Laparosc Endosc Percutan Tech. 2010;20(2):e60-3.

9. Kandil E, Noureldine S, Abdel Khalek M, Alrasheedi S, Aslam R, Friedlander P, Holsinger FC, Bellows CF. Initial experience using robot- assisted transaxillary thyroidectomy for Graves' disease. J Visc Surg. 2011;148(6):e447-51.

10. Inabnet WB 3rd. Robotic thyroidectomy: must we drive a luxury sedan to arrive at our destination safely? Thyroid. 2012;22(10):988-90.

11. Lin S, Chen ZH, Jiang HG, Yu JR. Robotic thyroidectomy versus endoscopic thyroidectomy: a meta-analysis. World J Surg Oncol. 2012;10:239.

12. Jackson NR, Yao L, Tufano RP, Kandil EH. Safety of robotic thyroidectomy approaches: Meta-analysis and systematic review. Head Neck. 2013 Mar 8. doi: 10.1002/hed.23223.

13. 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(3):850-61. doi: 10.1245/s10434-013-3406-7

14. Yi O, Yoon JH, Lee YM, Sung TY, Chung KW, Kim TY, Kim WB, Shong YK, Ryu JS, Hong SJ. Technical and Oncologic Safety of Robotic Thyroid Surgery. Ann Surg Oncol. 2013;20(6):1927-33.

15. Lee KE, Koo do H, Im HJ, Park SK, Choi JY, Paeng JC, Chung JK, Oh SK, Youn YK. Surgical completeness of bilateral axillo-breast approach robotic thyroidectomy: comparison with conventional open thyroidectomy after propensity score matching. Surgery. 2011;150(6):1266-74

16. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009;21;6(7):e1000097.

17. Wells GA, Shea B, O’Connell D, Peterson J, Welch V, Losos M et al. The Newcastle–Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp [accessed 2 August 2014].

18. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003;327(7414):557-60. Review

19. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315(7109):629-34.

20. Lee J, Nah KY, Kim RM, Ahn YH, Soh EY, Chung WY. Differences in postoperative outcomes, function, and cosmesis: open versus robotic thyroidectomy. Surg Endosc. 2010;24(12):3186-94.

21. Landry CS, Grubbs EG, Morris GS, Turner NS, Holsinger FC, Lee JE, Perrier ND. Robot assisted transaxillary surgery (RATS) for the removal of thyroid and parathyroid glands. Surgery. 2011;149(4):549-55.

22. Lee S, Ryu HR, Park JH, Kim KH, Kang SW, Jeong JJ, Nam KH, Chung WY, Park CS. Early surgical outcomes comparison between robotic and conventional open thyroid surgery for papillary thyroid microcarcinoma. Surgery. 2012;151(5):724-30.

23. 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(5):617-25.

24. Aliyev S, Taskin HE, Agcaoglu O, Aksoy E, Milas M, Siperstein A, Berber E. Robotic transaxillary total thyroidectomy through a single axillary incision. Surgery. 2013;153(5):705-10.

25. Ryu HR, Lee J, Park JH, Kang SW, Jeong JJ, Hong JY, Chung WY. A Comparison of Postoperative Pain After Conventional Open Thyroidectomy and Transaxillary Single-Incision Robotic Thyroidectomy: A Prospective Study. Ann Surg Oncol. 2013;20(7):2279-84

26. Kim WW, Kim JS, Hur SM, Kim SH, Lee SK, Choi JH, Kim S, Lee JE, Kim JH, Nam SJ, Yang JH, Choe JH. Is robotic surgery superior to endoscopic and open surgeries in thyroid cancer? World J Surg. 2011;35(4):779-84.

27. Lee S, Lee CR, Lee SC, Park S, Kim HY, Son H, Kang SW, Jeong JJ, Nam KH, Chung WY, Park CS, Cho A. Surgical completeness of robotic thyroidectomy: a prospective

comparison with conventional open thyroidectomy in papillary thyroid carcinoma patients. Surg Endosc. 2013 Nov 13. doi: 10.1007/s00464-013-3303-y

28. Kang SW, Jeong JJ, Yun JS, Sung TY, Lee SC, Lee YS, Nam KH, Chang HS, Chung WY, Park CS. Robot-assisted endoscopic surgery for thyroid cancer: experience with the first 100 patients. Surg Endosc. 2009;23(11):2399-406

29. Lee J, Kang SW, Jung JJ, Choi UJ, Yun JH, Nam KH, Soh EY, Chung WY. Multicenter study of robotic thyroidectomy: short-term postoperative outcomes and surgeon ergonomic considerations. Ann Surg Oncol. 2011;18(9):2538-47.

30. Tae K, Ji YB, Jeong JH, Lee SH, Jeong MA, Park CW. Robotic thyroidectomy by a gasless unilateral axillo-breast or axillary approach: our early experiences. Surg Endosc. 2011;25(1):221-8.

31. Broome JT, Pomeroy S, Solorzano CC. Expense of robotic thyroidectomy: a cost analysis at a single institution. Arch Surg. 2012;147(12):1102-6.

32. Cabot JC, Lee CR, Brunaud L, Kleiman DA, Chung WY, Fahey TJ 3rd, Zarnegar R. Robotic and endoscopic transaxillary thyroidectomies may be cost prohibitive when compared to standard cervical thyroidectomy: a cost analysis. Surgery. 2012;152(6):1016-24.

33. Foley CS, Agcaoglu O, Siperstein AE, Berber E. Robotic transaxillary endocrine surgery: a comparison with conventional open technique. Surg Endosc. 2012;26(8):2259-66.

34. Lee J, Na KY, Kim RM, Oh Y, Lee JH, Lee J, Lee JS, Kim CH, Soh EY, Chung WY. Postoperative functional voice changes after conventional open or robotic thyroidectomy: a prospective trial. Ann Surg Oncol. 2012;19(9):2963-70.

35. Tae K, Kim KY, Yun BR, Ji YB, Park CW, Kim DS, Kim TW. Functional voice and swallowing outcomes after robotic thyroidectomy by a gasless unilateral axillo-breast approach: comparison with open thyroidectomy. Surg Endosc. 2012;26(7):1871-7.

36. Yoo JY, Chae YJ, Cho HB, Park KH, Kim JS, Lee SY. Comparison of the incidence of postoperative nausea and vomiting between women undergoing open or robot-assisted thyroidectomy. Surg Endosc. 2013;27(4):1321-5.

37. Kim JA, Kim JS, Chang MS, Yoo YK, Kim DK. Influence of carbon dioxide insufflation of the neck on intraocular pressure during robot-assisted endoscopic thyroidectomy: a comparison with open thyroidectomy. Surg Endosc. 2013;27(5):1587-93.

Table 1. The ten articles which were excluded after reviewing the full-length text

First Author	Journal	Publication year, country	Title	Main reason for exclusion
Kang[28]	Surgical Endoscopy	2009, Korea	Robot-assisted endoscopic surgery for thyroid cancer: experience with the first 100 patients	Data from this study were included in a later study[22]
Lee[29]	Annals of Surgical Oncology	2011, Korea	Multicenter study of robotic thyroidectomy: short-term postoperative outcomes and surgeon ergonomic considerations.	There was no open thyroidectomy group for comparison.
Tae[30]	Surgical Endoscopy	2011, Korea	Robotic thyroidectomy by a gasless unilateral axillo-breast or axillary approach: our early experience	Data from this study were included in a later study[23]
Broome[31]	Archives of Surgery	2012, USA	Expense of robotic thyroidectomy: a cost analysis at a single institution	This study did not compare outcomes between robotic and open approaches
Cabot[32]	Surgery	2012, USA	Robotic and endoscopic transaxillary thyroidectomies may be prohibitive when compared to standard cervical thyroidectomy: a cost analysis	This study did not compare outcomes between robotic and open approaches
Foley[33]	Surgical Endoscopy	2012, USA	Robotic transaxillary endocrine surgery: a comparison with conventional open technique	Data from this study were included in a later study[24]
Lee[34]	Annals of Surgical Oncology	2012, Korea	Postoperative functional voice changes after conventional open or robotic	Data from this study were included in an earlier but

			thyroidectomy: a prospective trial	more representative study[19]
Tae[35]	Surgical Endoscopy	2012, Korea	Functional voice and swallowing outcomes after robotic thyroidectomy by a gasless unilateral axillo-breast approach: comparison with open thyroidectomy	Data from this study were included in a later study[23]
Yoo[36]	Surgical Endoscopy	2013, Korea	Comparison of the incidence of postoperative nausea and vomiting between women undergoing open or robot-assisted thyroidectomy	No oncological outcomes were reported
Kim[37]	Surgical Endoscopy	2013, Korea	Influence of carbon dioxide insufflation of the neck on intraocular pressure during robot-assisted endoscopic thyroidectomy: a comparison with open thyroidectomy.	No oncological outcomes were reported

Table 2a. A comparison of patient characteristics between robotic assisted thyroidectomy (RT) and open thyroidectomy (OT). Studies were grouped according to robotic approaches.

First author (year)	**Type**	**Number of patients**				**Mean age (±SD) (yrs)**		**Sex ratio (Male:Female)**		**Histological type**		**Study quality#**
		RT		**OT**		**RT**	**OT**	**RT**	**OT**	**RT**	**OT**	
		TT	**LTT**	**TT**	**LTT**							
Robotic transaxillary approach (RTAA)												
Lee (2010) [20]	RS	26	15	26	17	39.0 ± 7.0	37.7 ± 6.5	3:38	3:40	PTC = 41 FTC = 0	PTC = 43 FTC = 0	3
Landry (2012) [21]	RS	0	4	0	3	*50 (22 -62)	*53 (24 -75)	0:4	0:3	PTC = 3 FTC = 1	PTC = 2 FTC = 1	2
Lee (2012) [22]	RS	27	165	90	176	41.9 ± 9.2	48.7 ± 10.8	13:179	53:213	PTC = 192 FTC = 0	PTC = 266 FTC = 0	4
Tae (2012) [23]	RS	29	46	204	22	39.6 ± 8.9	51.0 ± 12.5	5:70	37:189	PTC = 75 FTC = 0	PTC = 226 FTC = 0	3
Aliyev (2013)	RS	11	0	21	0	48 ± 4	51 ± 3	0:11	1:20	PTC = 10 FTC = 1	PTC = 20 FTC = 1	2

[24]												
Ryu (2013) [25]	RS	45	0	45	0	39.0 ± 7.8	48.9 ± 10.3	3:42	9:36	PTC = 45 FTC = 0	PTC = 45 OT = 0	3
Yi (2013) [14]	RS	98	0	423	0	42.2 ± 8.2	51.8 ± 10.5	0:98	0:423	PTC = 98 FTC = 0	PTC = 423 FTC = 0	4
Lee (2013) [27]	PS	43	0	51	0	39.8 ± 10.2	48.3 ± 10.6	NR	NR	PTC = 43 FTC = 0	PTC = 51 FTC = 0	4
RTAA overall	-	279	230	860	218	41.1 ± 8.7	50.0 ± 10.8	24:427	102:901	PTC = 507 FTC =2	PTC= 1076 FTC = 2	
Robotic bilateral axillo-breast approach (RBA)												
Kim (2011) [26]	PS	69	0	138	0	41.3 ± 7.8	51.8 ± 8.9	6:63	34:104	PTC = 69 FTC = 0	PTC= 138 FTC = 0	4
Lee (2011) [15]	RS	174	0	237	0	39.9 ± 8.8	51.1 ± 11.1	20:154	50:187	PTC = 174 FTC = 0	PTC = 237 FTC = 0	4
RBA subtotal	-	243	0	375	0	40.3 ± 8.5	51.4 ± 10.3	26:217	84:291	PTC = 243 FTC = 0	PTC = 375 FTC = 0	
Overall	-	522	230	1235	218	40.8 ±	50.3 ±	50:644	186:11	PTC = 750	PTC= 1451	

						8.6	10.7		92	FTC = 2	FTC = 2	

Matching: 1 = age; 2 = sex; 3 = body mass index (BMI); 4 = histology; 5 = extent of thyroidectomy

*only the median and range were provided

Abbreviations: PS = prospective study; RS = retrospective study; NR = not reported; TT = total thyroidectomy; LTT = less than total thyroidectomy; PTC = papillary thyroid carcinoma; FTC = follicular thyroid carcinoma

using the Newcastle–Ottawa Scale

Table 2b. A comparison of tumor characteristics between robotic assisted thyroidectomy (RT) and open thyroidectomy (OT). Studies were grouped according to robotic approaches.

First author (year)	**Extra-thyroidal (%)**		**Multi-centricity (%)**		**Bilaterality (%)+**		**Mean tumor size (mm)**		**Number of CLNs retrieved (extent of CND)**		**CLN metastasis (%)**		**7th edition *TNM* stage (I/II/III/IV)**	
	RT	**OT**	**RT**	**OT**	**RT**	**OT**	**RT**	**OT**	**RT**	**OT**	**RT**	**OT**	**RT**	**OT**
Robotic transaxillary approach (RTAA)														
Lee (2010) [20]	NR	NR	10 (24.4)	9 (20.9)	3 (11.5)	5 (19.2)	8.3 ± 2.7	8.9 ± 3.0	4.4 ± 2.1 (UL)	4.3 ± 2.9 (UL)	18 (43.9)	19 (44.2)	37/0/4/0	38/0/5/0
Landry (2012) [21]	NR	NR	NR	NR	NR	NR	6.6 ± 7.7	7.0 ± 7.2	NR (no CND)	NR (no CND)	NR	NR	4/0/0/0	3/0/0/0
Lee (2012) [22]	86 (44.8)	131 (49.2)	47 (24.5)	79 (29.7)	22 (81.5)	49 (54.4)	6 ± 2.1	6 ± 2.1	4.6 ± 3.5 (UL)	5.7 ± 4.3 (UL)	49 (25.5)	66 (24.8)	150/0/42/0	167/0/96/3
Tae (2012) [23]	29 (38.6)	93 (41.2)	10 (13.3)	72 (31.9)	2 (6.9)	51 (25.0)	8.1 ± 4.7	9.9 ± 6.4	4.3 ± 2.4 (UL)	5.5 ±3.3 (UL)	26/57 (45.6)	64/153 (41.8)	60/0/15/0	142/5/77/2
Aliyev (2013)	NR	NR	NR	NR	NR	NR	18 ± 3	18 ± 2	NR (no CND)	NR (no CND)	NR	NR	9/1/1/0	17/1/3/0

[24]														
Ryu (2013) [25]	NR	NR	3 (6.7)	11 (24.4)	13 (28.9)	18 (40.0)	9.6 ± 3.6	11.8 ± 6.2	5.7 ± 4.8 (UL)	7.0 ± 5.2 (UL)	13 (28.9)	16 (35.6)	NR	NR
Yi (2013) [14]	63 (64.3)	267 (63.1)	34 (34.7)	164 (38.8)	24 (24.5)	97 (22.9)	*8 (1 – 19)	*8 (1 – 20)	*6.5 (0 – 4) (UL)	*7.0 (0 – 28) (UL)	36 (36.7)	160 (37.8)	67/0/ 31/0	183/0 /240/ 0
Lee (2013) [27]	39 (90.7)	41 (80.4)	18 (41.9)	15 (29.4)	15 (34.9)	13 (25.5)	10.1 ± 4.6	10.6 ± 7.1	4.9 ± 2.9 (UL)	6.3 ± 4.2 (UL)	17 (39.5)	27 (52.9)	31/0/ 12/0	22/0/ 29/0
RTAA subtotal	217 (53.2)	532 (55.1)	122 (24.7)	350 (33.2)	79 (29.5)	233 (27.8)	7.9 ± 3.3	8.8 ± 4.8	4.7 ± 3.3#	5.7 ± 4.0#	159 (33.4)	352 (35.9)	331/1 /132/ 0	560/6 /462/ 5
Robotic bilateral axillo-breast approach (RBA)														
Kim (2011) [26]	NR	NR	NR	NR	NR	NR	6 ± 2	7 ± 2	4.7 ± 2.7 (UL)	4.8 ± 2.8 (UL)	NR	NR	NR	NR
Lee (2011) [15]	131 (75.3)	178 (75.1)	78 (44.8)	88 (37.1)	42 (24.1)	43 (18.1)	8 ± 3.6	9 ± 4.9	NR (UL)	NR (UL)	82 (47.1)	97 (40.9)	123/0 /51/0	78/1/ 158/0
RBA	131	178	78	88	42	43	7.4 ±	8.3 ±	4.7 ±	4.8 ±	82	97	123/0	78/1/

subtotal	(75.3)	(75.1)	(44.8)	(37.1)	(24.1)	(18.1)	3.2	4.1	2.7#	2.8#	(47.1)	(40.9)	/51/0	158/0
Total	348 (59.8)	710 (59.0)	200 (29.9)	438 (33.9)	121 (27.4)	276 (25.7)	7.7 ± 3.3	8.6 ± 4.6	4.7 ± 3.2#	5.5 ± 3.8#	241 (37.1)	449 (36.9)	454/1/183/0	638/7/620/5
***p*-value**	0.732		0.368		0.721		<0.001		<0.001		0.745		0.001	

* only the median and range were available

only unilateral CND was analyzed

+ percentages were calculated based on total number of total thyroidectomies

Abbreviation: NR = not reported; CND = central neck dissection; UL = unilateral

Table 3. A comparison of postoperative stimulated thyroglobulin levels and locoregional recurrence between robotic-assisted thyroidectomy (RT) and open thyroidectomy (OT). Studies were grouped according to robotic approaches

First author (year)	Number of carcinoma		RAI ablation (%)+		Pre-ablation sTg level (ng/ml)		Post-therapy scan findings		LRR (%)		Mean follow-up (months)	
	RT	OT	RT	OT	RT	OT	RT	OT	RT	OT	RT	OT
Robotic trans-axillary approach (RTAA)												
Lee (2012) [22]^	192	266	NR	NR	0.25	0.22	No abnormal uptake	No abnormal uptake	0 (0.0)	0 (0.0)	29.1	29.1
Tae (2012) [23]	75	226	20 (69.0)	159 (77.9)	12.7± 15.0	4.9 ± 8.6	NR	NR	0 (0.0)	0 (0.0)	11.2	12.5
Aliyev (2013) [24]	11	21	4 (36.4)	11 (52.4)	1.0 ± 0.5*	1.0 ± 0.7*	0.65 ± 0.30##	0.84 ± 0.20##	0 (0.0)	0 (0.0)	9.0	9.0
Yi (2012) [14]^	98	423	88 (89.8)	363 (85.8)	2.1 ± 3.8	1.1 ± 4.2	NR	NR	0 (0.0)	1 (0.2)	27.4	28.8
Lee (2013) [27]	43	51	43 (100)	51 (100)	4.9 ± 1.4	4.2 ± 1.2	0.10 ± 0.01#	0.05 ± 0.01#	0 (0.0)	0 (0.0)	12	12

RTAA subtotal	419	987	155 (85.6)	584 (83.5)	6.1 ± 9.0	2.5 ± 5.8	0.15 ± 0.08	0.19 ± 0.08	0 (0.0)	1 (0.2)	17.7 ± 9.7	18.3 ± 9.8
Robotic bilateral axillo-breast approach (RBA)												
Kim (2011) [26]	69	138	NR	NR	0.8 ± 1.4	0.8 ± 2.0	NR	NR	NR	NR	NR	NR
Lee (2011) [15]^	174	237	174 (100)	237 (100)	1.4 ± 3.9	1.2 ± 3.1	12.8 ± 13.3#	13.5 ± 13.3#	0 (0.0)	0 (0.0)	17.3	19.9
RBA subtotal	243	375	174 (100.0)	237 (100.0)	1.2 ± 3.4	1.1 ± 2.7	12.8 ± 13.3#	13.5 ± 13.3#	0 (0.0)	0 (0.0)	17.3	19.9
Total	662	1362	329 (92.7)	821 (87.7)	3.6 ± 6.7	2.0 ± 5.0	10.1 ± 11.8	10.7 ± 11.9	0 (0.0)	1 (0.2)	17.7 ± 8.7	18.6 ± 8.8
***p*-value**	-		0.849		0.033		0.453		-		0.860	

Abbreviations: RAI = radioactive iodine; LRR = locoregional recurrence; sTg = stimulated thyroglobulin; NR = not reported

+ percentages were calculated based on total number of total thyroidectomies

*mean non-stimulated levels; # thyroid bed-to-background ratio;## percentage uptake of RAI

^studies which had been verified with the corresponding author

Figure legends

Figure 1. A flow diagram for study selection

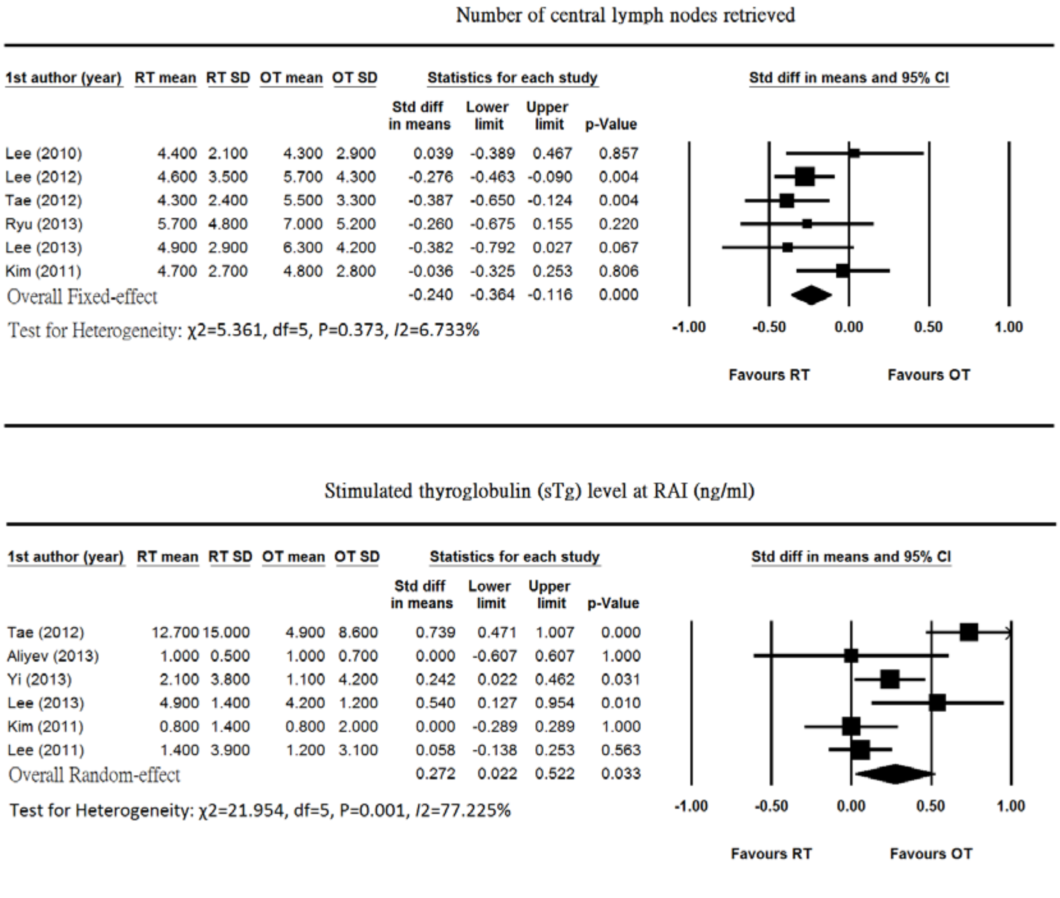

Figure 2a. A forest plot for number of central lymph nodes retrieved in robotic thyroidectomy (RT) and open thyroidectomy (OT); Figure 2b. A forest plot for pre-ablation stimulated thyroglobulin (sTg) level in robotic total thyroidectomy (RT) and open thyroidectomy (OT)