Original Article

A systematic review and meta-analysis comparing outcomes between robotic-assisted thyroidectomy and non-robotic endoscopic thyroidectomy

Running title: Robotic has better outcomes than non-robotic

Subject category: Clinical Research; Oncology/Endocrine

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

Carlos KH WONG[2], PhD

Julian Shun TSANG[1], MBBS, MRCS

Kai Pun WONG[1], MBBS, MRCS

[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

Address for Correspondence:

Dr Brian HH Lang

Department of Surgery

Queen Mary Hospital, 102 Pokfulam Road, Hong Kong SAR, China

Tel.: (852) 22554232; Fax No.: (852) 28172291

Email: blang@hkucc.hku.hk

Authors' contribution:

BHH Lang / CKH Wong / JS Tsang / KP Wong 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.

ABSTRACT

Background: Despite its feasibility, using the *da Vinci* robot in remote-access thyroidectomy remains controversial. This meta-analysis compared surgical and oncological outcomes between robotic-assisted thyroidectomy (RT) and non-robotic endoscopic thyroidectomy (ET).

Methods: A systematic review was performed to identify studies comparing outcomes between RT and ET. Outcomes included operating time, drain output, complications, number of central lymph nodes (CLNs) retrieved and pre-ablation stimulated thyroglobulin (sTg) level. A random-effects model was used.

Results: Six studies were eligible. Of the 3510 patients, 2167 (61.7%) underwent RT while 1343 (38.3%) underwent ET. Despite a higher drain output (185.8mls vs. 173.3 mls, p=0.019), RT had fewer temporary recurrent laryngeal nerve (RLN) injury (2.6% vs. 3.3%, p=0.035) and shorter LOS (3.4 days vs. 3.5 days, p=0.030). In terms of oncological outcomes, despite higher incidence of multicentricity and larger tumors, the number of CLNs retrieved during unilateral CND in RT was significantly greater than ET (4.5 ± 2.6 and 3.4 ± 2.5, p<0.001) while the pre-ablation sTg was comparable (0.8ng/mL vs. 1.1ng/mL, p=0.456). However, follow-up data were relatively scarce.

Conclusions

Adding the robot in remote-access thyroidectomy was associated with a significantly lower risk of temporary RLN injury and shorter LOS. However, despite achieving a comparable level of surgical completeness for low-risk differentiated thyroid carcinoma between RT and ET, this

study highlighted the limitations with the current literature and the need for more prospective studies with adequate follow-up.

Keywords: endoscopic thyroidectomy; non-robotic thyroidectomy; robotic thyroidectomy; total thyroidectomy; central neck dissection; papillary thyroid carcinoma; hypoparathyroidism; recurrent laryngeal nerve; nerve monitoring

INTRODUCTION

Thyroidectomy is a common surgical procedure associated with a low morbidity in experienced hands.[1] However, to further improve cosmetic result and patient satisfaction, various endoscopic approaches to the thyroid gland (or endoscopic thyroidectomy (ET)) have been developed.[2,3] Unlike the conventional approach, these approaches involve making incisions away from the neck (i.e. remote-access thyroidectomy) and they are generally technically challenging.[2,3] Furthermore, they involve working in a small working space with rigid endoscopic instruments.[2,3] In 2007, a South Korean group aimed to improve the ergonomics of ET by pioneering the use of the da Vinci robot (the so-called "robotic-assisted thyroidectomy" or RT).[4-6] Despite the higher cost, it has the advantages of improved flexibility of endoscopic instruments, availability of a more stable, three-dimensional view and lessening physiologic tremors.[4] However, despite these benefits, it remains unclear whether the addition of the robot in ET would translate into better outcomes.[7] Some studies found adding the robot prolonged the procedural time while other outcomes appeared comparable.[8-10] However, given the generally low incidence of surgically-related complications, these studies lacked the statistical power to detect a difference. To our knowledge, two meta-analyses have been published with one reporting comparable outcomes while the other reporting increased complications and drain output in the RT group.[11,12] Given the growing number of publications on this controversial subject and their indication has been extended to managing low-risk differentiated thyroid carcinoma (DTC), we conducted a systematic review and meta-analysis to compare the surgically-related complications and oncological outcomes between RT and ET.

METHODS

This systematic review and meta-analysis was conducted in accordance with the PRISMA statement.[13]

Search strategy

Studies comparing surgical-related and / or oncological outcomes between patients who underwent RT and ET were retrieved from the Scopus, Medline (PubMed) and Cochrane Library electronic databases on 12th November 2013. We used the following free text search terms in "All fields"

#1: 'robotic thyroidectomy'

#2: 'robotic assisted thyroidectomy'

#3: 'robot thyroidectomy'

#4: 'endoscopic thyroidectomy'

#5: #1 OR #2 OR #3 OR #4

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 independently screened 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. Since there were no randomized trials, any prospective or retrospective study comparing at least one surgically-related and / or oncological outcome between RT and ET was included. However, we excluded case reports, editorials, expert opinions, reviews without original data, studies on pediatric population, studies comparing outcomes between RT and open thyroidectomy and studies evaluating patients undergoing robotic-assisted lateral neck dissection. Surgically-related outcomes included operating time, postoperative drain output, recurrent laryngeal nerve (RLN) injury, hypoparathyroidism after total thyroidectomy (TT), hematoma formation, length of hospital stay (LOS), cosmetic result and any other possible complications such as infection, seroma, tracheal injury, chyle leakage, pain, brachial plexus injury and flap paresthesia. Oncological outcomes included number of central lymph nodes (CLNs) harvested during central neck dissection (CND), postoperative stimulated thyroglobulin (sTg) level and locoregional recurrence (LR). Multiple reports of the same dataset were assessed and the most representative and 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, selection method for RT and ET, weight / size of excised thyroid gland, number of lobectomies, number and extent of bilateral resections (TT or less-than total thyroidectomy (LTT)), pathology, characteristics of DTC, number of CLNs harvested during unilateral (i.e. not bilateral) CND, operating time, volume of drain output, rate and definition of surgically-related complications, radioiodine (RAI) ablation, postoperative sTg level and LR. TT included near-TT, TT and TT with CND whereas LTT only included subtotal thyroidectomy or Dunhill procedure. Operating time was the duration in minutes from skin incision to closure. Operating times were stratified according to the extent of resection (lobectomy, bilateral thyroid resections, LTT and TT). For studies that separately provided times for TT and LTT, a pooled estimate of the two was used to calculate overall mean in bilateral thyroid resection. Hypocalcemia rate was calculated by dividing the total number of hypocalcemia over the total number of TTs. RLN injury rate was calculated by dividing the total number of injuries over the total number of nerves at risk. In TT, two RLNs were considered at risk whereas in lobectomy and LTT, only one RLN was considered at risk.

Statistical analysis

For comparison of dichotomous variables between RT and ET, chi-square tests and Fisher's exact tests were used. Student t-test was used for comparison of continuous variables. The Pearson's correlation test was used to correlate two 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 total operating time, volume of drain output, LOS, tumor size, number of CLNs retrieved and postoperative sTg level and odds ratios (OR) were examined for the other surgical outcomes. Results were aggregated and analyzed using a random-effect model. Publication bias was estimated by Begg's rank correlation test and Egger's regression test.[14,15] 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 1750 titles initially identified from the database search, 11 full-length articles were assessed for inclusion, of which 5 were excluded and 6 studies[8,10,16-19] were determined to be eligible and were included in this systematic review. Table 1 lists these 5 articles[5,9,20-22] 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 5 articles excluded, 2 were excluded mainly because some of the data were superseded by one later study[17]. One study was excluded because outcomes / complications were not clearly defined[9].

Patient selection

Ultrasonography was used as a routine preoperative imaging modality in all studies[8,10,16-19]. The inclusion and exclusion criteria were similar in all studies. Inclusions included patient age between 21–65 years old, malignant tumor size ≤ 2–4 cm, benign nodule ≤5cm and body mass index ≤36kg/m^2.[8,10,16,18] One study only included papillary microcarcinoma.[19] Exclusions included previous neck irradiation, previous neck surgery, presence of lateral lymph node and distant metastases, severe thyroiditis, Graves' disease and posteriorly located carcinoma.[8,10,16,18,19] In terms of case selection, 2 studies[8,10] were based on patient preference while the other 4 studies[16-19] did not specify their selection method. One study[18] used historical ET controls for comparing outcomes with RT.

Baseline characteristics

Table 2 shows a comparison of the baseline characteristics between the 6 eligible studies. All 6 studies were retrospective. Of the 3510 patients included, 2167 (61.7%) had RT (RT group) while 1343 (38.3%) had ET (ET group). In the RT group, there were 901 (41.6%) TTs and 1266

(58.4%) LTTs while in the ET group, there were 310 (23.1%) TTs and 1033 (76.9%) LTTs. Overall, the RT group had greater proportion of TT than ET (p<0.001). The trans-axillary approach (TAA) was utilized in 5 studies[8,10,16-18] while the bilateral axillo-breast approach (BABA) was utilized in one study[19]. There was no comparison between RT and ET using different approaches (i.e. TAA vs. BABA).

Age and sex ratio were matched in all 6 studies[8,10,16-19]. The mean age was comparable (39.4 ± 8.8 vs. 38.2 ± 9.3 years old, p=0.243) but the proportion of females in RT was significantly less (92.9% vs. 97.8%, p<0.001). Weight / size of excised gland was compared in only 1 study and that study[8] found the weight of excised gland in ET was significantly heavier than RT. In terms of thyroid pathology, 4 studies[8,10,16,18] managed both benign and malignant diseases while 2 studies[17,19] managed malignant disease only.

Surgical outcomes

Table 3a shows a comparison of outcomes between the two groups. The overall mean drain output in RT was significantly greater (185.8 ± 93.2 and 173.3 ± 105.6mls, respectively) (SMD = 0.207, 95%CI: 0.034 to 0.379, p=0.019). Quantitative meta-analysis confirmed that the ET group had an overall mean reduced drain output 16.3 (95%CI= 0.4 – 32.1) mls. The potential publication bias did not appear significant, as confirmed by the Begg analysis (Kendall's tau = -0.500, p=0.308) and the Egger regression test (z= 1.289, p=0.326). Table 3b shows a comparison of operating times. For lobectomies, the mean operating time was not significantly different between RT and ET (156.1 ± 38.8 vs. 139.7 ± 49.4 mins) (SMD = 0.335, 95%CI: -0.799 to 1.468, p=0.563). For bilateral thyroid resections, the mean operating time in RT was also not significantly different from that in ET (135.3 ± 35.6 and 136.8 ±45.6 mins, respectively) (SMD =

0.240, 95%CI: -0.432 to 0.911, *p*=0.484). When only TT was considered, the mean operating time was also similar between the two groups (SMD = 0.295, 95%CI:-0.864 to 1.454, *p*=0.617)

Table 4 shows the definition used for postoperative hypocalcemia and RLN injury between the 6 studies. Four studies defined permanent hypocalcemia as the failure to have postoperative parathyroid hormone and / or adjusted serum calcium normalized within 6 months.[8,10,18,19] All 6 studies[8,10,16-19] reported their temporary postoperative hypocalcemia rate while 6 studies[8,10,16-19] reported permanent postoperative hypocalcemia rate. Assuming these studies adopted a similar definition of hypocalcemia, the overall temporary hypocalcaemia rate in the RT group was comparable to that of ET group (321/901 (35.6%) and 97/310 (31.3%) respectively; OR = 1.040, 95%CI=0.655 – 1.652, *p*=0.868) while the overall permanent hypocalcaemia was also comparable (1/901 (0.1%) and 5/310 (1.6%), respectively; OR=0.172, 95%CI=0.017 – 1.687, *p*=0.131).

Similar to hypocalcemia, the definition for temporary and permanent RLN injury varied between studies. All six studies[8,10,16-19] defined permanent RLN injury as persistent hoarseness or impairment in vocal cord function for > 6-month. Routine perioperative direct laryngoscopy (DL) was performed in 4 studies[8,10,16,18] while selective DL was reported in 2 other studies[17,19]. Figure 2 shows the forest plot for temporary RLN injury. The cumulative temporary RLN injury rate in RT was significantly lower than ET (79/3068 (2.6%) and 54/1653 (3.3%), respectively) (OR=0.681; 95%CI= 0.476 – 0.973, *p*=0.035). Both the Begg (Kendall's tau = 0.400, *p*=0.260) and Egger regression analyses (z= 0.416, *p*=0.699) did not reveal any potential publication bias. To see if temporary RLN injury was volume related, the number of RT cases reported was correlated with temporary RLN injury rate. However, there was no significant correlation between the number of RT cases reported and the temporary RLN injury rate observed (ρ =

0.422, *p*=0.405). The cumulative permanent RLN palsy was comparable between RT and ET groups (10/3068 (0.3%) and 5/1653 (0.3%), respectively) (OR=1.153, 95%CI= 0.267 – 4.980, *p*=0.848).

Hematoma rate was reported in all 6 studies[8,10,16-19]. The cumulative hematoma rate was comparable between RT and ET (13/2167 (0.6%) and 14/1343 (1.0%), respectively) (OR=0.550, 95%CI=0.255 – 1.185, *p*=0.127). Also the rate of other surgically-related complications (including seroma, wound infection, skin burn, tracheal injury, brachial plexus injury and chyle leak) was similar between RT and ET (70/2167 (3.2%) and 39/1343 (2.9%), respectively) (OR=1.017, 95%CI=0.682 to 1.518, *p*=0.933). On the other hand, LOS in RT was significantly shorter than ET (3.4 ± 1.2 and 3.5 ± 1.2 days, respectively) (SMD=-0.079, 95%CI=-0.149 to -0.008, *p*=0.030). Potential publication bias did not appear significant, as confirmed by the Begg analysis (Kendall's tau = 0.067, p=0.188) and the Egger regression test (z= 0.076, p=0.943). Figure 3 shows the forest plot for LOS.

Other reported outcomes

Altogether 3 studies compared postoperative pain and of these, one study[18] found comparable pain score on first day and first week between the two groups while one study[8] found significantly higher pain score in RT on postoperative day 0. However, this study also found one less surgical assistant required in RT than in ET.[8] As expected, the incidence of brachial plexus injury and skin flap paraesthesia appeared similar between the two groups.[16,17] Cosmetic satisfaction also appeared similar.[18] However, the learning curve for RT appeared to be significantly shorter than for ET.[16,17] When measured by the number of cases required before reaching a plateau in operating time, RT required 35 to 45 cases while ET required 55 – 60 cases.[16,17] One study compared the direct medical cost of the two procedures and found RT to be at least 8.0 times more costly than ET.[10]

Tumor characteristics and oncological outcomes

Table 5 shows a comparison of tumor characteristics. Only one study compared the proportion of extrathyroidal extension and there was no significant difference between RT and ET (p=0.132). On the other hand, there were significant differences between RT and ET in tumor multi-centricity (511/2013 (25.4%) vs. 121/944 (12.8%), OR=2.319, 95%CI=1.869 – 2.878, p<0.001), bilaterality (217/2013 (10.8%) vs. 72/944 (7.6%), OR=1.456, 95%CI=1.102 – 1.925, p=0.008) and primary tumor size (5.5 ± 5.1mm and 5.1 ± 4.8mm, respectively) (SMD=0.167; 95%CI=0.093 – 0.240, p<0.001). Among the 6 studies reporting the number of CLNs retrieved during surgery, five studies[8,10,16,17,19] performed unilateral CND while one study[18] performed a combination of unilateral and bilateral CND. When only unilateral CND was considered, the overall mean number of CLNs retrieved in RT was not significantly different from that of the ET group (4.5 ± 2.6 and 3.4 ± 2.5, respectively) (SMD=0.402, 95%CI=-0.040 to 0.844, p=0.075). Table 6 shows a comparison of RAI ablation, postoperative sTg levels and mean follow-up between the two groups. Of the 2 studies[16,17] which provided the overall percentage of RAI ablation given, none provided separate percentage for RT and ET. Ablation doses ranged from 30 – 150 mCi and were followed by ^{131}whole body scan (WBS) 5 – 7 days later.[16-18] None of the 531 patients in the two reported studies had any abnormal uptake on the WBS. Four studies[10,16,17,19] compared the pre-ablation sTg level between the two groups. The overall mean pre-ablation sTg in RT was comparable to ET (0.8 ± 1.9ng/mL and 1.1 ± 3.4ng/mL, respectively) (SMD=-0.024, 95%CI: -0.183 to 0.136, p=0.772). Although 4 studies provided the number of

LR in the two groups,[10,16-18] only one provided the mean follow-up[16] and so it was not possible to compare the overall incidence rate ratio between the two groups.

DISCUSSION

To our knowledge, this is the largest and most comprehensive meta-analyses aimed at comparing the outcomes between RT and ET. In contrast to previous meta-analyses,[11,12] our analysis showed that RT was associated with a significantly reduced temporary RLN injury (2.6% vs. 3.3%, *p*=0.035) and shorter LOS (3.4 days vs. 3.5 days, *p*=0.030). In concordance with one previous meta-analysis,[11] our analysis found that the RT group had significantly higher drain output than ET (*p*=0.003). However, this might be attributed to the greater proportion of TT:LTT in the RT group. Also because of this, when we calculated the RLN injury rate, the number of nerves at risk was used instead of the number of patients. In fact, when it was calculated based on number of patients, the difference in temporary RLN injury was not significant (*data not shown*). Since previous meta-analyses did not account for this, we postulate that this was why they did not find a significantly lower temporary RLN injury rate in the RT group. Another reason might have been because there was an addition of a large multi-center study[17] which was not available at the time of the two previous meta-analyses[11,12].

In terms of outcomes, there were no significant differences in total operating time when the robot was used and this appeared to be irrespective of the number of lobes removed (i.e. lobectomy or bilateral thyroid resection), extent of thyroidectomy (lobectomy or TT) and surgical approach (i.e. TAA or BABA). Although this finding might seem surprising as one might expect more time required in RT because of the extra time required for docking and undocking of the robot, we believe operating time in RT could be significantly reduced through better surgical manipulations and endoscopic view. Furthermore, unlike ET which normally uses three ports (i.e. one for camera and two for working instruments), RT normally uses four arms and that might have helped to further shorten the time for dissection. Also when the operating time was stratified into stages, the actual console time (i.e. the time when the surgeon was actually operating) in experienced hands could be as short as 50 minutes while the docking time as short as 5 mins.[20] Similarly, the lower temporary RLN injury rate in RT might have been related to improved surgical manipulations and more stable endoscopic view. However, since routine DL was only used in 4 studies[8,10,16,18] and asymptomatic vocal cord palsy may be missed without routine DL,[23] the actual injury rate could have been higher. Nevertheless, since similar protocol was used for both groups in all 6 studies, routine DL probably may not have changed the significant difference in temporary RLN injury between the two groups. Regarding the LOS, given that the extent of surgery and the drain output were significantly greater in the RT group, we were surprised that it was significantly shorter in the RT (3.4 vs. 3.5 days, p=0.003). However, given such small difference and other factors could have potentially affected LOS, further studies are necessary to confirm this finding. Similar to previous meta-analyses,[11,12] other outcomes such as temporary (p=0.868) and permanent hypocalcemia (p=0.131), hematoma (p=0.127) and the sum of other complications (including seroma, tracheal injury, brachial injury, chyle leak ..etc) (p=0.933) appeared comparable between the two groups.

Although higher direct medical cost is an obvious disadvantage with RT, a recent cost analysis has found that when both the endoscopic equipment depreciation cost and complications cost had been added together, the actual cost difference between RT and ET (USD 13670 vs. USD 12505, respectively) was significantly less than what had been reported previously.[9,10,22] The other potential advantage of RT was the shorter learning curve than ET which is relevant for a surgeon who would like to learn either procedure.[16,17]

Unlike surgically-related outcomes, oncological outcomes were more difficult to assess and compare. This was because firstly, it appeared that most of the tumor characteristics such as multi-centricity, bilaterality and tumor size were not well-matched between the two groups. Secondly, data on postoperative follow-up were relatively scarce and too short. Only 1 of 6 studies provided the mean follow-up duration in patients with DTC. Also given that most were microcarcinoma with good tumor risk, a larger cohort with significantly longer follow-up would be necessary to truly evaluate their long-term outcomes. Nevertheless, with these limited data, both groups had comparable postoperative sTg level (0.8ng/mL vs. 1.1ng/mL, p=0.456) and so we would conclude that for low risk DTC, the oncological outcome after RT appeared equivalent to that after ET.

Despite these findings, our data should be interpreted cautiously because all 6 eligible studies were non-randomized and retrospective and so their findings were subjected to selection biases. These biases were partly reflected by the differences in baseline and tumor characteristics and these could potentially have affected both the surgically-related and oncological outcomes observed. Also outcomes such as complications varied between studies. There were two different approaches used for matching which made pooled data difficult to interpret. Furthermore, the effect of learning curve was not accounted for. For example, in many institutions including ours,[8] ET preceded RT and so some of the data from ET might actually represent the earlier part of the learning curve and this tended to favor RT (such as fewer RLN injury and shorter LOS). In addition, we would like to acknowledge that due to the study design, we had to exclude one previous study[20] which only contained some (but not all) patients in a latter study[17]. Therefore, future prospective studies are necessary to confirm our findings.

Conclusion

RT was associated with a significantly lower risk of temporary RLN injury and shorter LOS than ET while other complications appeared comparable. Although RT appeared to have comparable surgical completeness and short-term oncological outcomes as ET for low-risk DTC, there were too few studies with adequate follow-up data to allow a definite conclusion. Future prospective studies with longer follow-up are required to confirm our findings.

AUTHOR DISCLOSURES

All authors including Dr LANG Hung-Hin Brian, Dr Carlos KH WONG, Dr Julian Shun TSANG and Dr Kai Pun WONG have nothing to disclose. We declared that we have had no corporate or commercial relationships that might pose a conflict of interest.

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 July 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;44:185-198.

3. Lang BH, Lo CY. Technological innovations in surgical approach for thyroid cancer. J Oncol. 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;22:986-987.

5. 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; 23:2399-406.

6. 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:e1-7.

7. Lang BH, Chow MP, Wong KP. Endoscopic vs robotic thyroidectomy: which is better? Ann Surg Oncol. 2011;18 Suppl 3:S251.

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

9. Kiriakopoulos A, Linos D. Gasless transaxillary robotic versus endoscopic thyroidectomy: exploring the frontiers of scarless thyroidectomy through a preliminary comparison study. Surg Endosc 2012;26:2797-801.

10. Yoo H, Chae BJ, Park HS, Kim KH, Kim SH, Song BJ, Jung SS, Bae JS. Comparison of surgical outcomes between endoscopic and robotic thyroidectomy. J Surg Oncol 2012;105:705-8.

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. 2014 ;36:137-43.

13. 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.

14. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Biometrics 1994;50(4):1088-101.

15. 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.

16. Lee J, Lee JH, Nah KY, Soh EY, Chung WY. Comparison of endoscopic and robotic thyroidectomy. Ann Surg Oncol 2011;18:1439-46

17. Lee J, Yun JH, Choi UJ, Kang SW, Jeong JJ, Chung WY. Robotic versus Endoscopic Thyroidectomy for Thyroid Cancers: A Multi-Institutional Analysis of Early Postoperative Outcomes and Surgical Learning Curves. J Oncol 2012;doi: 10.1155/2012/734541.

18. Tae K, Ji YB, Jeong JH, Kim KR, Choi WH, Ahn YH. Comparative study of robotic versus endoscopic thyroidectomy by a gasless unilateral axillo-breast or axillary approach. Head Neck 2013; 35:477-84.

19. 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; 35:779-84.

20. Lee S, Ryu HR, Park JH, Kim KH, Kang SW, Jeong JJ, Nam KH, Chung WY, Park CS. Excellence in robotic thyroid surgery: a comparative study of robot-assisted versus conventional endoscopic thyroidectomy in papillary thyroid microcarcinoma patients. Ann Surg 2011;253(6):1060-6.

21. 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;18:2538-47.

22. 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.

23. Wong KP, Lang BH, Ng SH, Cheung CY, Chan CT, Lo CY. A prospective assessor-blind evaluation of surgeon-performed transcutaneous laryngeal ultrasonography in vocal cord function before and after thyroidectomy. Surgery 2013 ;154:1158-64
24. Lang BH, Lo CY, Chan WF, Lam KY, Wan KY. Prognostic factors in papillary and follicular thyroid carcinoma: implications for cancer staging. Ann Surg Oncol 2007; 14:730-8

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

First Author	**Journal**	**Publication year, country**	**Title**	**Main reason for exclusion**
Kang[5]	Surgical Endoscopy	2009, Korea	Robot-assisted endoscopic surgery for thyroid cancer: experience with the first 100 patients	There was no endoscopic thyroidectomy group for comparison
Lee[20]	Annals of Surgery	2011, Korea	Excellence in robotic thyroid surgery. A comparative study of robot-assisted versus conventional endoscopic thyroidectomy in papillary thyroid microcarcinoma patients	Some data from this study were included in a later study[17]
Lee[21]	Annals of Surgical Oncology	2011, Korea	Multicenter study of robotic thyroidectomy: short-term postoperative outcomes and surgeon ergonomic considerations	This study only compared surgeon's musculoskeletal discomfort and ergonomics
Cabot[22]	Surgery	2012, USA	Robotic and endoscopic transaxillary thyroidectomies may be prohibitive when compared to standard cervical thyroidectomy: a cost analysis	Data from this study were included in a later study[17]
Kiriakopoulos[9]	Surgical Endoscopy	2012, Greece	Gasless transaxillary robotic versus endoscopic thyroidectomy: exploring the frontiers of scarless thyroidectomy through a preliminary comparison study.	Complications were not clearly defined

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

First author (year)	Number of patients				Mean age (±SD) (yrs)		Sex ratio (Male:Female)		Weight / size of gland	Final pathology		Match between RT and ET
	RT		ET		RT	ET	RT	ET		Benign	Malign.	
	TT	LTT	TT	LTT								
Trans-axillary approach (TAA)												
Lang (2011)[8]	4	3	10	29	42.1 ± 8.2	41.3 ± 7.9	0:7	1:38	ET>RT	RT=6 ET=35	RT=1 ET=4	1,2,5,6
Lee (2011)[16]	48	115	2	94	38.7 ± 8.2	39.9 ± 6.5	6:157	2:94	NR	RT=11 ET=41	RT=152 ET=55	1,2,3
Lee (2012)[17]	706	1063	150	693	39.4 ± 9.1	37.5 ± 9.4	132:1637	19:824	NR	RT=0 ET=0	RT=1769 ET=843	1,2,5
Yoo (2012)[10]	30	16	49	116	37.4 ± 8.1	38.9 ± 9.1	0:46	0:165	NR	RT=2 ET=17	RT=44 ET=148	1,2,3,5
Tae (2013)[18]	44	69	4	101	40.0 ±	38.9 ±	10:103	6:99	NR	RT=21	RT=92	1,2,3

					9.3	10.6				ET=59	ET=46	
TAA overall	**832**	**1266**	**215**	**1033**	**39.3 ± 8.9**	**38.1 ± 9.4**	**148:1950**	**28:1220**	**-**	**RT=40 ET=152**	**RT=2058 ET=1096**	**-**
Bilateral Axillo-Breast Approach (BABA)												
Kim (2011)[19]	69	0	95	0	41.3 ± 7.8	39.9 ± 9.1	6:63	2:93	NR	RT=0 ET=0	RT=69 ET=95	1,2,3,5,6
BABA overall	69	0	95	0	41.3 ± 7.8	39.9 ± 9.1	6:63	2:93	NR	RT=0 ET=0	RT=69 ET=95	-
Overall	**901**	**1266**	**310**	**1033**	**39.4 ± 8.8**	**38.2 ± 9.3**	**154:2013**	**30:1313**	**-**	**RT=40 ET=152**	**RT=2127 ET=1191**	**-**

Matching: 1 = age; 2 = sex; 3 = body mass index (BMI); 4 = weight of excised thyroid gland; 5 = final pathology; 6 = extent of thyroidectomy

Abbreviations: NR = not reported; TT = total thyroidectomy; LTT = less than total thyroidectomy

Table 3a. A comparison of surgical outcomes between robotic-assisted thyroidectomy (RT) and endoscopic thyroidectomy (ET). Studies were grouped according to robotic approaches.

First author (year)	**Drain output (mls)**	**Hypocalcemia# (%)**		**RLN injury (%)+**		**Hematoma (%)**	**Sum of other complications (%)***	**Other outcomes**	**LOS (days)**
		Temp	**Perm**	**Temp**	**Perm**				
Lang (2011)[8]	NR	RT=0 (0.0) ET=0 (0.0)	RT=0 (0.0) ET=0 (0.0)	RT=0 (0.0) ET=3 (6.1)	RT=1 (9.1) ET=0 (0.0)	RT=0 (0.0) ET=0 (0.0)	RT=0 (0.0) ET=0 (0.0)	Pain, number of surgical assistants	RT=2.0 ± 1.0 ET=2.0 ± 1.0
Lee (2011)[16]	RT=152.8 ± 90.7 ET=119.8 ± 75.3	RT=6 (12.5) ET=0 (0.0)	RT=0 (0.0) ET=0 (0.0)	RT=3 (1.4) ET=3 (3.1)	RT=1 (0.5) ET=1 (1.0)	RT=1 (0.6) ET=3 (3.1)	RT=7 (4.3) ET=3 (3.1)	Learning curve	RT=2.8 ± 1.1 ET=3.2 ± 1.9
Lee (2012)[17]	NR	RT=276 (39.1)	RT=0 (0.0)	RT=68 (2.7)	RT=8 (0.3)	RT=10 (0.6) ET=8 (1.0)	RT=52 (2.9) OT=27 (3.2)	Learning curve	RT=3.3 ± 1.3

		ET=55 (36.7)	ET=2 (0.2)	ET=41 (4.1)	ET=1 (0.1)				ET=3.4 ± 1.1
Yoo (2012)[10]	RT=188.1 ± 71.2 ET=190.3 ± 60.54	RT=5 (16.7) ET=18 (36.7)	RT=0 (0.0) ET=0 (0.0)	RT=0 (0.0) ET=1 (0.5)	RT=0 (0.0) ET=0 (0.0)	RT=0 (0.0) ET=0 (0.0)	RT=1 (2.2) ET=0 (0.0)	Cost	RT=2.9 ± 1.1 ET=3.0 ± 1.2
Tae (2013)[18]	RT=254.9 ± 116.2 ET=231.7 ± 187.6	RT=11 (25.0) ET=0 (0.0)	RT=0 (0.0) ET=0 (0.0)	RT=7 (4.5) ET=4 (3.7)	RT=0 (0.0) ET=1 (0.9)	RT=2 (1.8) ET=2 (1.9)	RT=8 (7.1) ET=8 (7.6)	Pain, cosmetic satisfaction	RT=5.9 ± 0.6 ET=5.8 ± 1.3
TAA overall	**RT=193.7 ± 98.2 ET=183.7 ± 115.0**	**RT=298 (35.8) ET=73 (34.0)**	**RT=0 (0.0) ET=2 (0.9)**	**RT=78 (2.7) ET=52 (3.6)**	**RT=10 (0.4) ET=3 (0.2)**	**RT=13 (0.6) ET=13 (1.0)**	**RT= 68 (3.2) ET= 38 (3.0)**	**-**	**RT=3.4 ± 1.3 ET=3.5 ± 1.2**
Kim (2011)[19]	RT=149.0 ± 64.8	RT=23 (33.3)	RT=1 (1.4)	RT=1 (1.4)	RT=0 (0.0)	RT=0 (0.0) ET=1 (1.1)	RT=2 (2.9) ET=1 (1.1)	-	RT=3.1 ± 0.7

	ET=133.1 ± 56.2	ET=24 (25.3)	ET=3 (3.2)	ET=2 (2.1)	ET=2 (2.1)				ET=3.1 ± 0.9
Overall	**RT=185.8 ± 93.2 ET=173.3 ± 105.6**	**RT=321 (35.6) ET=97 (31.3)**	**RT=1 (0.1) ET=5 (1.6)**	**RT=79 (2.6) ET=54 (3.3)**	**RT=10 (0.3) ET=5 (0.3)**	**RT=13 (0.6) ET=14 (1.0)**	**RT= 70 (3.2) ET= 39 (2.9)**	**-**	**RT=3.4 ± 1.2 ET=3.5 ± 1.2**

Abbreviations: NR = not reported; RLN = recurrent laryngeal nerve; temp = temporary; perm = permanent; LOS = length of stay

only total thyroidectomy was analyzed

*including seroma, tracheal injury, wound infection, skin burn, brachial plexus injury and chyle leakage within that group

+calculated based on number of nerves-at-risk

Table 3b. A comparison of operating times between robotic-assisted thyroidectomy (RT) and endoscopic thyroidectomy (ET). Studies were grouped according to robotic approaches

	Operating time (in minutes)		
	Lobectomy	**Bilateral thyroid resection**	
		LTT	**TT**
Trans-axillary approach (TAA)			
Lang (2011)[8]	RT=102.5 +/- 30 ET=96 +/- 28	-	RT=161 +/- 32.9 ET=135 +/- 31.0
Lee (2011)[16]	-	RT=110 +/- 50.7 ET=142.7 +/- 52.1	
Lee (2012)[17]*	-	RT=122.3 +/- 32.4 ET=127.2 +/- 41.3	RT=149.2 +/- 32.3 ET=172.7 +/- 66.7
Yoo (2012)[10]	RT=161.0 +/- 44.5 ET=118.3 +/- 36.0	-	RT=189.6 +/- 71.2 ET=145.2 +/- 35.5
Tae (2013)[18]	RT=157.3 +/- 37.6 ET=177.6 +/- 65.2	-	RT=184.5 +/- 42.3 ET=241.2 +/- 46.6
TAA overall	**RT=156.1 ± 38.8** **ET=139.7 ± 49.4**	**RT=133.2 ± 35.3** **ET=136.9 ± 46.7**	
Bilateral Axillo-Breast Approach (BABA)			
Kim (2011)[19]	-	-	RT=196 +/- 45.0 ET=136 +/- 31.0
BABA overall	**-**	**-**	**RT=196 +/- 45.0** **ET=136 +/- 31.0**
Overall	**RT=156.1 ± 38.8** **ET=139.7 ± 49.4**	**RT=135.3 ± 35.6** **ET=136.8 ± 45.6**	

Abbreviations: LTT = less than total thyroidectomy; TT = total thyroidectomy

*the pooled estimates for bilateral thyroid resection were RT=133.0±32.4 and ET=135.3±46.8

Table 4. Definitions of postoperative hypocalcemia and recurrent laryngeal nerve injury in the six eligible studies

First author (year)	Temporary hypocalcaemia	Permanent hypocalcaemia	Preoperative laryngoscopy	Temporary RLN injury	Permanent RLN injury	Postoperative laryngoscopy
Lang (2011)[8]	Symptomatic or serum adjusted Ca < 2mmol/L ≤6 months	Symptomatic or serum adjusted Ca < 2mmol/L for > 6 months	Routine	Vocal cord palsy recovered within 6 months	Vocal cord palsy failed to recover after 6 months	Routine. One week after operation
Kim (2011)[19]	Postoperative PTH normalized within 6 months	Postoperative PTH failed to normalize after 6 months	Selective, based on symptoms	Hoarseness and/or vocal cord palsy lasting ≤ 6 months after surgery	Hoarseness and/or vocal cord palsy > 6 months after surgery	Selective, based on symptoms
Lee (2011)[16]	Not defined	Not defined	Routine	Vocal cord palsy recovered within 6 months	Vocal cord palsy failed to recover after 6 months	Routine. One week after operation
Lee (2012)[17]	Not defined	Not defined	Selective, based on symptoms	Transient hoarseness, resolved within 6 months	Hoarseness, failed to resolve after 6 months	Selective, based on symptoms
Yoo (2012)[10]	Symptomatic hypocalcemia or PTH <5pg/mL for ≤6 months	Symptomatic hypocalcemia or PTH <5pg/mL > 6 months	Routine	Recovery of vocal cord palsy within 6 months after surgery	Vocal cord palsy for > 6 months after surgery	Routine

Tae (2013)[18]	PTH below normal for ≤6 months	PTH below normal for >6 months	Routine	Vocal cord palsy recovered within 6 months	Vocal cord palsy failed to recover after 6 months	Routine.

Abbreviations: RLN = recurrent laryngeal nerve; PTH = parathyroid hormone

Table 5. A comparison of tumor characteristics between robotic assisted thyroidectomy (RAT) and open thyroidectomy (ET). Studies were grouped according to robotic approaches.

First author (year)	Extra-thyroidal (%)		Multi-centricity (%)		Bilaterality (%)		Mean (±SD) tumor size (mm)		Number of central lymph nodes retrieved#		Matching between RT and ET
	RT	ET	RT	ET	RT	ET	RT	ET	RT	ET	
Trans-axillary approach (TAA)											
Lang (2011)[8]	NR	NR	NR	NR	NR	NR	11 ± 0.0	17 ± 2.0	6.0 ± 0.0	4.7 ± 2.0	4,5
Lee (2011)[16]	NR	NR	29 (19.1)	8 (14.5)	5 (3.3)	0 (0.0)	8.7 ± 6.4	8.4 ± 4.1	4.5 ± 1.5	2.4 ± 1.9	2,3,4,6
Lee (2012)[17]	NR	NR	469 (26.5)	110 (13.0)	208 (11.8)	72 (8.5)	5 ± 5	4 ± 5	4.5 ± 2.6	2.9 ± 1.7	2,3,4
Yoo (2012)[10]	NR	NR	NR	NR	NR	NR	8.7 ± 6.4	8.4 ± 4.1	4.7 ± 3.5	5.4 ± 4.5	4,5,6
Tae	39	10	13	3 (6.5)	4 (4.3)	0 (0.0)	8.6 ± 5.9	8.4 ± 7.6	4.7 ± 2.9	4.3 ± 3.6	1,2,3,4,5

(2013)[18]	(42.4)	(21.7)	(14.1)								
TAA overall	**39 (42.4)**	**10 (21.7)**	**511 (25.4)**	**121 (12.8)**	**217 (10.8)**	**72 (7.6)**	**5.5 ±5.2**	**5.0 ± 5.0**	**4.5 ± 2.6**	**3.3 ± 2.4**	**-**
Bilateral Axillo-Breast Approach (BABA)											
Kim (2011)[19]	NR	NR	NR	NR	NR	NR	6 ± 2	6 ± 2	4.7 ± 2.7	4.6 ± 3.7	4,5
BABA overall	NR	NR	NR	NR	NR	NR	6 ± 2	6 ± 2	4.7 ± 2.7	4.6 ± 3.7	-
Overall	**39 (42.4)**	**10 (21.7)**	**511 (25.4)**	**121 (12.8)**	**217 (10.8)**	**72 (7.6)**	**5.5 ± 5.1**	**5.1 ± 4.8**	**4.5 ± 2.6**	**3.4 ± 2.5**	**-**

Matching: 1 = extrathyroidal extension; 2 = multicentricity; 3 = bilaterality; 4 = tumor size; 5 = number of central lymph nodes retrieved; 6 = tumor stage

#only unilateral central neck dissection was analyzed

NR = not reported

Table 6. A comparison of postoperative stimulated thyroglobulin levels and locoregional recurrence between robotic-assisted thyroidectomy (RT) and endoscopic thyroidectomy (ET). Studies were grouped according to robotic approaches

First author (year)	Number of carcinoma		RAI ablation (%)		Postoperative sTg level (ng/ml)		Number of LR (%)		Follow-up (months)	
	RT	ET	RT	ET	RT	ET	RT	ET	RT	ET
Trans-axillary approach (TAA)										
Lee (2011)[16]	152	55	NR#	NR#	1.0 ± 0.7	1.0 ± 4.7	0 (0.0)	0 (0.0)	12	12
Lee (2012)[17]	1769	843	NR#	NR#	0.7 ± 1.8	0.6 ± 2.0	3 (0.2)	3 (0.4)	NR	NR
Yoo (2012)[10]^	44	148	NR	NR	3.8 ± 5.2	3.5 ± 5.6	0 (0.0)	0 (0.0)	NR	NR
Tae (2013)[18]	92	46	NR	2 (50.0)	NR	NR	0 (0.0)	0 (0.0)	NR	NR
Bilateral Axillo-Breast Approach (BABA)										
Kim (2011)[19]	69	95	NR	NR	0.8 ± 1.4	2.4 ± 6.3	NR	NR	NR	NR

Abbreviations: sTg = stimulated thyroglobulin; NR = not reported

#did not provide separate data for RT and ET

^ verified with the corresponding author

LEGENDS

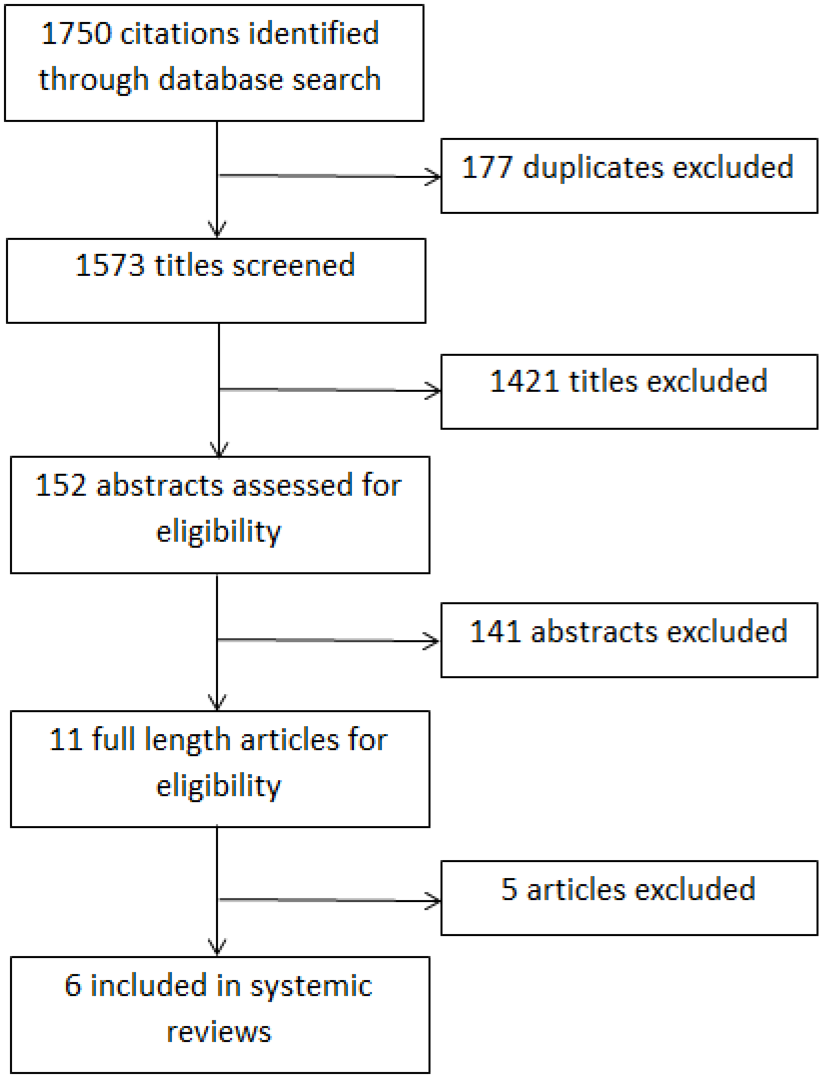

Figure 1. A flow diagram for study selection

Temporary recurrent laryngeal nerve injury

1st author & year	Temp RLN Injury / Total: RT	ET	Statistics for each study: Odds ratio	Lower limit	Upper limit	p-Value
Lang (2011)	0 / 11	3 / 49	0.578	0.028	11.988	0.723
Lee (2011)	3 / 211	3 / 98	0.457	0.091	2.305	0.343
Lee (2012)	68 / 2475	41 / 993	0.656	0.442	0.973	0.036
Yoo (2012)	0 / 76	1 / 214	0.930	0.037	23.080	0.965
Tae (2013)	7 / 157	4 / 109	1.225	0.350	4.291	0.751
Kim (2011)	1 / 138	2 / 190	0.686	0.062	7.643	0.759
Random-effect overall estimate			0.681	0.476	0.973	0.035

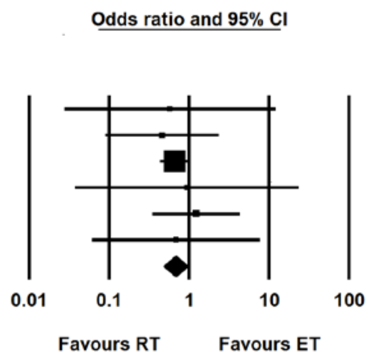

Figure 2. A forest plot for temporary recurrent laryngeal nerve injury

Hospital length of stay

1st author & year	RT mean	RT SD	ET mean	ET SD	Statistics for each study			
					Std diff in means	Lower limit	Upper limit	p-Value
Lang (2011)	2.0	1.0	2.0	1.0	0.000	-0.805	0.805	1.000
Lee (2011)	2.8	1.1	3.2	1.9	-0.276	-0.529	-0.023	0.033
Lee (2012)	3.3	1.3	3.4	1.1	-0.081	-0.163	0.001	0.054
Yoo (2012)	2.9	1.1	3.0	1.2	-0.085	-0.412	0.242	0.611
Tae (2013)	5.9	0.6	5.8	1.3	0.100	-0.166	0.366	0.461
Kim (2011)	3.1	0.7	3.1	0.9	0.000	-0.310	0.310	1.000
Random-effect overall estimate					-0.079	-0.149	-0.008	0.030

Std diff in means and 95% CI

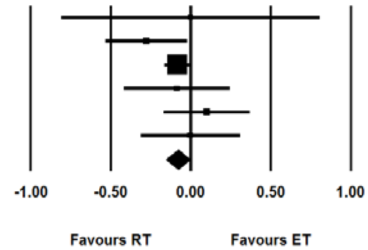

Figure 3. A forest plot for length of hospital stay