

Systematic Review and Meta-analysis of Robotic vs Conventional Thyroidectomy Approaches for Thyroid Disease

Gordon H. Sun, MS, MD^{1,2,3*}, Lilia Peress^{1*}, and Melissa A. Pynnonen, MSc, MD¹

Otolaryngology–
 Head and Neck Surgery
 1–13
 © American Academy of
 Otolaryngology—Head and Neck
 Surgery Foundation 2014
 Reprints and permission:
 sagepub.com/journalsPermissions.nav
 DOI: 10.1177/0194599814521779
 http://otojournal.org



Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

Abstract

Objective. This study compared postoperative technical, quality-of-life, and cost outcomes following either robotic or open thyroidectomy for thyroid nodules and cancer.

Data Sources. PubMed, Ovid MEDLINE, EMBASE, ISI Web of Science, and the Cochrane Central Register of Controlled Trials.

Review Methods. We examined relevant controlled trials, comparative effectiveness studies, and cohort studies for eligible publications. We calculated the pooled relative risk for key postoperative complications, mean differences for operative time, and standardized mean differences for length of stay (LOS) using random effects models. Quality-of-life outcomes were summarized in narrative form.

Results. The meta-analysis comprised 11 studies with 726 patients undergoing robotic transaxillary or axillo-breast thyroidectomy and 1205 undergoing open thyroidectomy. There were no eligible cost-related studies. Mean operative time for robotic thyroidectomy exceeded open thyroidectomy by 76.7 minutes, while no significant difference in LOS was identified. There were no significant differences in hematoma, seroma, recurrent laryngeal nerve injury, hypocalcemia, or chyle leak rates. The systematic review included 12 studies. Voice, swallowing, pain, and paresthesia outcomes showed no significant differences between the 2 approaches. The robotic cohort reported higher cosmetic satisfaction scores, although follow-up periods did not exceed 3 months and no validated questionnaires were used.

Conclusions. Transaxillary and axillo-breast robotic and open thyroidectomy demonstrate similar complication rates, but robotic approaches may introduce the risk of new complications and require longer operative times. Robotic thyroidectomy appears to improve cosmetic outcomes, although longer follow-up periods and use of validated instruments are needed to more rigorously examine this effect.

Keywords

robotic surgery, thyroid nodule, thyroid cancer, length of stay, operative time, quality of life, hemorrhage, recurrent

laryngeal nerve injury, hoarseness, hypoparathyroidism, hypocalcemia, brachial plexus injury

Received August 16, 2013; revised January 6, 2014; accepted January 9, 2014.

The traditional surgical approach for thyroidectomy, the transcervical incision popularized by Kocher, remained relatively unchanged from its inception in the late 19th century until 1997, when an endoscopic approach was described.¹⁻³ Subsequently, in 2000, the da Vinci surgical robot (Intuitive Surgical, Inc, Sunnyvale, California) was approved by the US Food and Drug Administration (FDA) for certain laparoscopic surgical procedures.⁴ The surgical robot allows the surgeon to perform the operation by manipulating the robot arms while seated at a control panel, without the need for the surgeon to hold and manipulate an endoscopic or other surgical instrument by hand, thus reducing tremor and improving fine motor control of instrumentation. Since then, the surgical robot has been used to develop minimally invasive approaches to thyroidectomy, with or without complementary endoscopy. However, the use of the surgical robot for thyroidectomy is not FDA approved and remains off-label.

There is substantial controversy over the role of the robot in thyroidectomy because traditional open thyroidectomy already has a low morbidity rate and produces excellent results. Improved cosmesis is the primary advantage of robotic thyroidectomy, which trades a cervical incision for one in the axillary space or anterior chest wall.^{5,6} Other

¹Department of Otolaryngology–Head and Neck Surgery, University of Michigan, Ann Arbor, Michigan, USA

²Robert Wood Johnson Foundation Clinical Scholars, University of Michigan, Ann Arbor, Michigan, USA

³VA Center for Clinical Management Research, Ann Arbor VA Healthcare System, Ann Arbor, Michigan, USA

*These authors contributed equally to this work.

Corresponding Author:

Gordon H. Sun, MS, MD, Partnership for Health Analytic Research, LLC, 280 South Beverly Dr, Suite 404, Beverly Hills, CA 90212, USA.
 Email: gsun@pharllc.com

advantages may include improved surgeon ergonomics and dexterity.⁷ Limitations of robotic thyroidectomy include longer operative time, increased equipment and staffing needs, and complications related to the surgical approach, such as brachial plexus injury.

A previous meta-analysis compared robotic, endoscopic, and conventional open thyroidectomy approaches in 9 studies, concluding that the robotic approach was safe and efficacious.⁸ However, the study examined only pain and cosmesis and did not study other important quality-of-life (QoL) outcomes such as voice and swallowing function. Another recent meta-analysis of 6 studies compared robotic and endoscopic thyroidectomies and concluded that the risk of complications was higher using the robotic approach, but the authors did not define the nature of the complications of interest.⁹

We conducted the following meta-analysis to examine postoperative complications and provide a more comprehensive overview of QoL-related issues in clinical trials and cohort studies comparing robotic and conventional open thyroidectomy for thyroid cancer and nodules. Our secondary objective was to characterize the methodological quality of the literature.

Patients and Methods

Search Strategy

The findings from this systematic review and meta-analysis were reported in accordance with the 2009 Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.¹⁰ The protocol was registered with the PROSPERO database (#CRD42012003331). On January 2, 2013, we searched PubMed, Ovid MEDLINE, EMBASE, ISI Web of Science, and the Cochrane Central Register of Controlled Trials. We also manually searched the bibliographies of published review articles and other relevant publications for additional studies. Due to limited resources, we did not review conference proceedings and abstracts, gray literature databases, and non-English publications. The full search strategy is available in the Appendix (available at otojournal.org) and at PROSPERO (http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42012003331). The complete search yielded 670 references. After duplicate entries were removed, 277 references remained. The references were imported into EndNote (Thomson Reuters, New York, New York) for abstract review. This study was exempt from review by the University of Michigan Medical School Institutional Review Board.

Study Selection

Two authors (G.H.S. and L.P.) screened the titles and abstracts of the 277 articles, with disagreements being resolved by consensus or arbitration with a third author (M.A.P.). We included randomized and nonrandomized controlled trials, comparative effectiveness studies, and prospective or retrospective cohort studies in which robotic and conventional open approaches to the thyroid were compared in patients with thyroid nodules or cancer. We included studies using endoscopic thyroidectomy

techniques only if the endoscope was used in the surgical robot arm and the comparison group was open thyroidectomy. According to our a priori protocol,¹¹ all eligible studies were required to report any one of the following: (1) at least 1 commonly accepted postthyroidectomy health-related outcome such as surgical complications, operative time, or hospital length of stay (LOS); (2) at least 1 QoL measure such as voice, swallowing, or cosmetic satisfaction; or (3) actual treatment costs for robotic and open thyroidectomy patients. We excluded studies for which a full-length manuscript was not available (eg, conference abstracts).

Data Extraction and Publication Quality Assessment

The same 2 authors independently abstracted the following variables: number of patients; patient age, sex, race/ethnicity, and country where the study took place; body mass index (BMI); benign vs malignant pathology; tumor size or extent of disease; and total vs partial thyroidectomy, neck dissection, and surgical approach. We recorded complications, robot-to-open conversion rate, operative time, LOS, and treatment costs. We evaluated publication quality according to the Martin criteria, which are 10 elements believed to be essential to allow comparison of surgical outcomes across practices and institutions.¹²

Systematic Review Approach and Statistical Analysis

We described patient demographic and tumor characteristics using descriptive statistics. We compared complication rates between groups using 2-sample *t* tests and tests of proportions. For meta-analysis, we calculated pooled relative risk (RR) with 95% confidence intervals (CIs) to summarize categorical complications. We compared operative times using mean differences with 95% CIs and hospital LOS with standardized mean differences (SMDs). The SMD is defined as the difference between the mean outcomes of the compared groups divided by the pooled standard deviation of the outcome. We used DerSimonian and Laird random effects models to pool outcomes across studies. We used forest plots to depict quantitative outcomes and Cochran's *Q* and *I*² tests to examine heterogeneity between studies. We conducted all statistical analyses with Stata 12.1 (StataCorp LP, College Station, Texas) and used 2-sided (*P* < .05) tests of significance.

Results

Study Characteristics

We reviewed the titles and abstracts and excluded 261 of the 277 articles (**Figure 1**). We reviewed full-length manuscripts of the remaining 16 publications and excluded 4 more studies. We excluded Broome et al¹³ because the costs for robotic surgery were not based on the authors' original data but instead extrapolated from 2 previously published studies. We excluded Cabot et al¹⁴ because the study reported outcomes selectively and Lee et al⁷ because no patient outcomes were reported. We excluded Yoo et al¹⁵ because the diagnoses were unclear (eg, thyroid nodule vs thyroiditis). Twelve publications were included in the

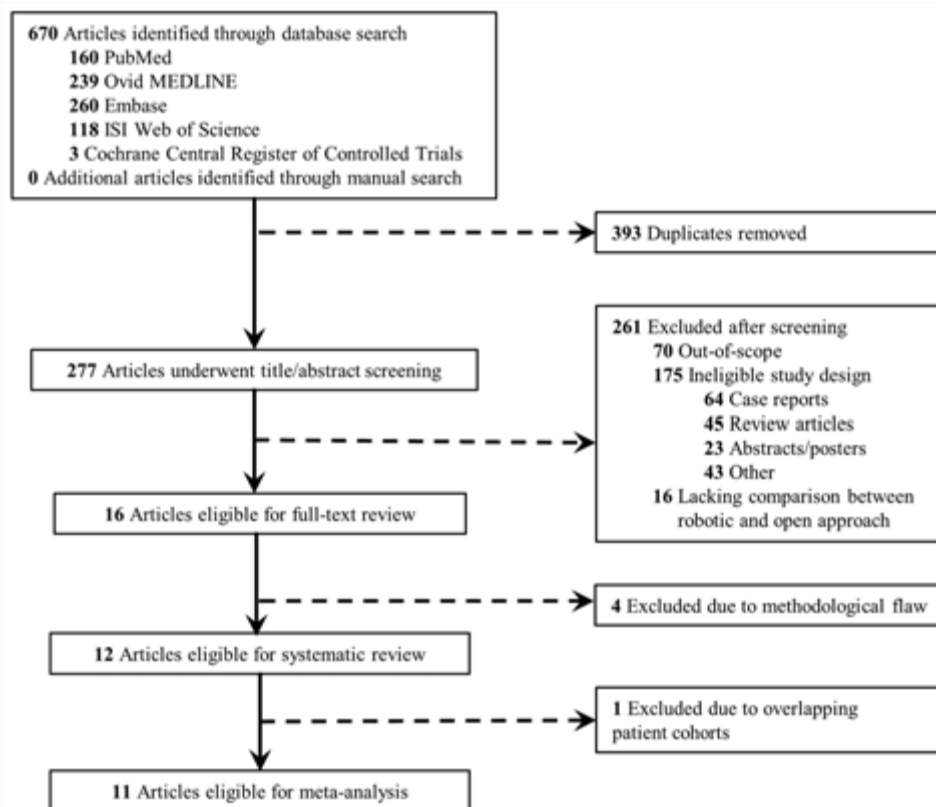


Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flowchart of study selection.

systematic review (**Table 1** and online Appendix). None of the articles described cost.

We contacted study authors to obtain or clarify data. From Dr Kyung Tae (Hanyang University, Seoul, South Korea),^{16,17} we obtained previously unpublished individual-level data spanning October 2008 to October 2010 (personal communication, May 25, 2013), which encompassed 2 published studies from his group that had overlapping cohorts.^{16,17} Altogether, there were 97 robotic and 307 open cases from Tae's group. We used these unpublished data rather than the individual published studies^{16,17} from his group to avoid case duplication. Another article from Tae's group¹⁸ was included in the systematic review but not the meta-analysis because it included only outcomes related to QoL.

The 11 articles in the meta-analysis reported characteristics and outcomes for 1931 patients: 726 patients undergoing robotic thyroidectomy and 1205 patients undergoing open thyroidectomy. **Table 2** compares the demographic and surgical characteristics of all patients by surgical approach. There were significant differences in age, sex, BMI, and the proportion undergoing total thyroidectomy. The robotic group was younger (40.5 vs 49.2 years), had more women (91.2% vs 78.5%), had a lower mean BMI (23.1 vs 24.2), and underwent fewer total thyroidectomies (58.1% vs 75.1%).

Methodological Quality and Assessment of Bias

Using Martin criteria, we analyzed 8 articles for publication quality (0-10, with higher scores indicating higher quality),

finding that the mean score of the studies was 7.75 (median, 8; range, 3-9).^{16,17,19-24} All studies described the method of patient accrual, morbidity rate, and complications; however, definitions of complications varied considerably. For example, Lee et al²³ and Tae et al¹⁶ defined permanent vocal cord paralysis as no evidence of recovery within 6 months, while Lee et al²⁵ used 3 months as the cutoff. Follow-up duration was reported in 5 of 8 studies.^{16,17,20,23,24} The longest period of follow-up was 29.1 months (range, 10-38 months),²⁰ while the other studies followed patients from 3 months up to 1 year.

Patient Outcomes

Operative time. Seven articles^{16,17,19-22,26} reported overall operative time, while 4 articles^{16,17,23,24} reported time separately for subtotal vs total thyroidectomies. Operative time was significantly longer in the robotic thyroidectomy group (**Figures 2-4**). The mean operative time of robotic thyroidectomy exceeded conventional open thyroidectomy by 76.7 minutes (95% confidence interval [CI], 47.3-106.1). There was significant heterogeneity among the studies ($I^2 = 95.3\%$, $P < .001$). There was an increased mean difference of 48.1 minutes (95% CI, 33.6-62.7) via robotic approaches specifically in total thyroidectomy patients and 37.3 minutes (95% CI, 12.0-62.5) via robotic approaches in subtotal thyroidectomy patients. There was significant heterogeneity noted in both of these comparisons ($I^2 = 84.8\%$, $P = .001$ for total thyroidectomy; $I^2 = 97.3\%$, $P < .001$ for subtotal thyroidectomy).

Table 1. Characteristics of All 12 Studies Incorporated into Systematic Review.

Study (Author, Year)	Study Type	Country	Total No. of Patients (Open, Robotic)	Robotic Approach	Follow-up Period
Foley et al, ¹⁹ 2012	Retrospective	United States	27 (16, 11)	Unilateral transaxillary, with endoscope, gasless	Perioperative
Kang et al, ²⁰ 2012	Retrospective	South Korea	165 (109, 56)	Unilateral transaxillary, with endoscope, gasless	Tg at 6 mo, US at 1 y
Kim et al, ²⁶ 2013	Prospective	South Korea	37 (18, 19)	Bilateral axillo-breast, with endoscope and insufflation	Perioperative
Kim et al, ²¹ 2011	Retrospective	South Korea	207 (138, 69)	Bilateral axillo-breast, with endoscope and insufflation	Perioperative
Landry et al, ²² 2012	Retrospective	United States	50 (25, 25)	Unilateral transaxillary, with endoscope; use of insufflation not addressed	Perioperative
Lee et al, ²⁵ 2012	Prospective	South Korea	88 (46, 42)	Unilateral transaxillary, without endoscope, gasless	1 wk, 3 mo
Lee et al, ²³ 2010	Prospective	South Korea	84 (43, 41)	Unilateral transaxillary, with endoscope, gasless	1 wk, 3 mo
Lee et al, 2011 ²⁷	Prospective	South Korea	411 (237, 174)	Bilateral axillo-breast, with endoscope and insufflation	Tg at 8-12 wk
Lee et al, ²⁴ 2012	Retrospective	South Korea	458 (266, 192)	Unilateral axillo-breast, with endoscope, gasless	Mean 29.1 mo (range, 10-38 mo)
Tae et al, ¹⁶ 2012 ^a	Retrospective	South Korea	301 (226, 75)	Unilateral axillo-breast or axillary, with endoscope, gasless	Mean 11.18 mo in robotic, 12.53 mo in open
Tae et al, ¹⁸ 2012 ^b	Prospective	South Korea	111 (61, 50)	Unilateral axillo-breast, with endoscope, gasless	1 d, 1 wk, 1 mo, 3 mo, 6 mo
Tae et al, ¹⁷ 2011 ^a	Retrospective	South Korea	208 (167, 41) ^c	Unilateral axillo-breast (2/41 axillary without breast port), with endoscope, gasless	1 wk, 1 mo, 3 mo

Abbreviations: Tg, thyroglobulin; US, ultrasound.

^aThese studies, plus additional information obtained directly from the corresponding author, were combined for meta-analysis.

^bThis was excluded from meta-analysis due to reporting of only quality-of-life outcomes.

^cThere were 167 open cases in the Methods section but only 163 in Table 1 of the Results section of the referenced article.

Table 2. Demographic and Surgical Characteristics of Patients from All Eligible Studies.^a

Characteristic	Robotic Cohort (n = 726)	Open Cohort (n = 1205)	P Value
Mean age, y	40.5	49.2	<.001
Female sex, No. (%)	662 (91.2)	947 (78.5)	<.001
Body mass index ^b	23.1	24.2	<.001
Thyroid malignancy, No. (%) ^c	660/707 (93.4)	1111/1187 (93.6)	.83
Total thyroidectomy, No. (%)	422 (58.1)	905 (75.1)	<.001
Tumor size, mm ^d	7.9	8.0	.64

^aNot all patient characteristics were reported in every study. $P < .05$ represents the threshold of statistical significance.

^bBased on 298 robotic and 434 open cases.

^cBased on 707 robotic and 1187 open cases.

^dBased on 682 robotic and 1162 open cases.

Hospital length of stay. Seven articles^{16,17,20-24} reported LOS. The SMD between open and robotic thyroidectomy hospital LOS was -0.006 (95% CI, -0.25 to 0.24), meaning that robotic thyroidectomy patients had a small,

nonsignificant lower mean LOS compared with their open thyroidectomy counterparts (**Figure 5**). Significant heterogeneity among studies was observed ($I^2 = 75.1\%$, $P = .001$).

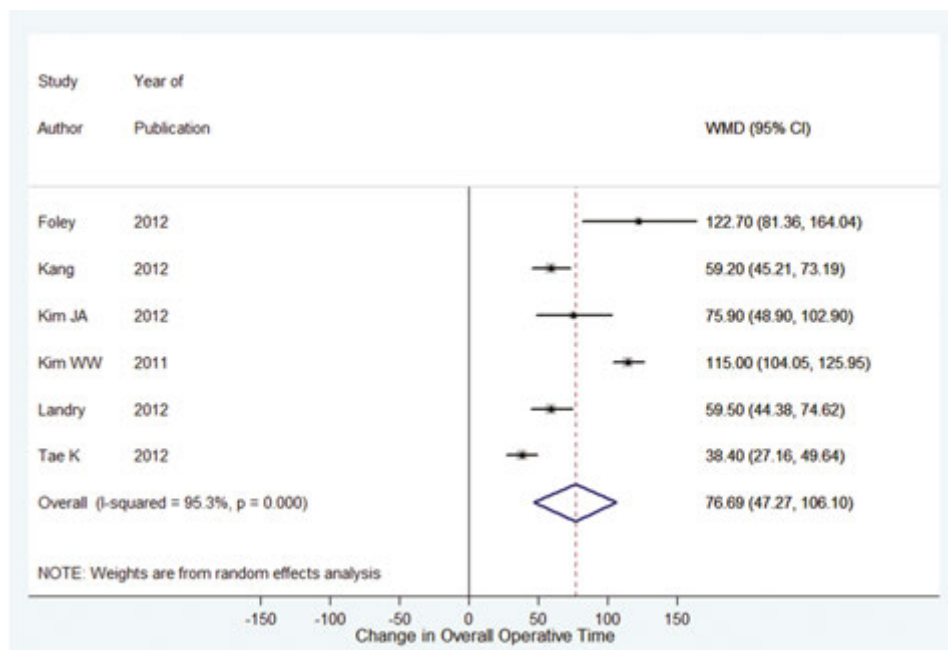


Figure 2. Total operative time across all types of thyroidectomies, robotic vs open approach. CI, confidence interval; WMD, weighted mean difference.

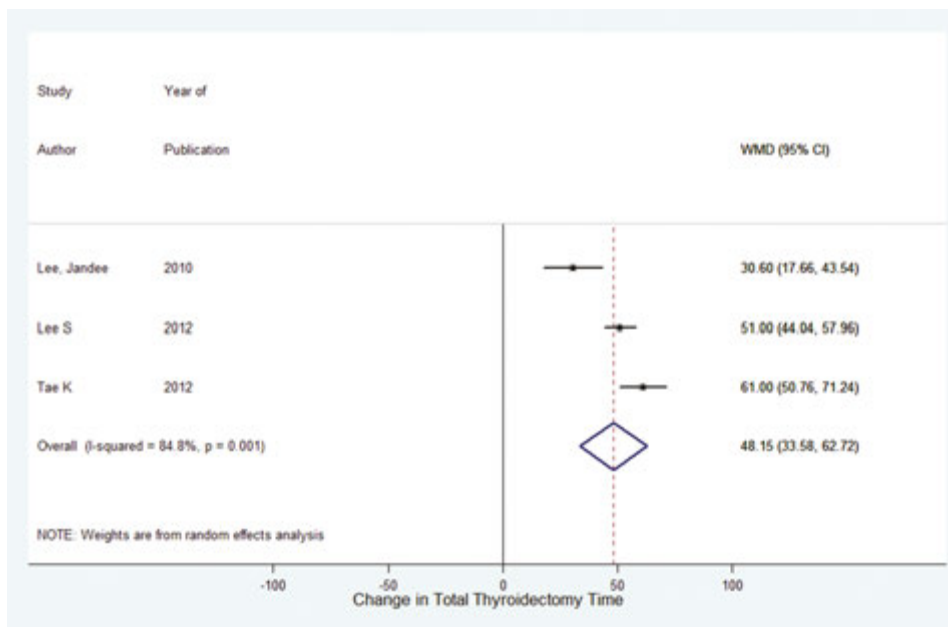


Figure 3. Total thyroidectomy operative time, robotic vs open approach. CI, confidence interval; WMD, weighted mean difference.

Postoperative complications. There were no statistically significant differences between robotic and open thyroidectomy outcomes in the rates of hematoma, seroma, recurrent laryngeal nerve (RLN) injury, hypocalcemia, or chyle leak (**Figures 6-12**). There was no significant heterogeneity among studies for any complication except temporary hypocalcemia ($I^2 = 83.7\%$, $P < .001$). There were no deaths or tracheal injuries reported in any of the studies.

Conversion from robotic to open thyroidectomy was necessary in 2 of 243 (0.8%) cases.^{17,19,21-23} The 6 articles^{16,17,19,22-24} that evaluated brachial plexus injury reported a combined event rate of 8 of 366 (2.2%). Foley et al¹⁹ reported 1 patient with discomfort on shoulder movement that resolved prior to the first postoperative visit. Landry et al²² reported 1 patient with a decreased range of motion due to pain and a second patient with trouble gripping

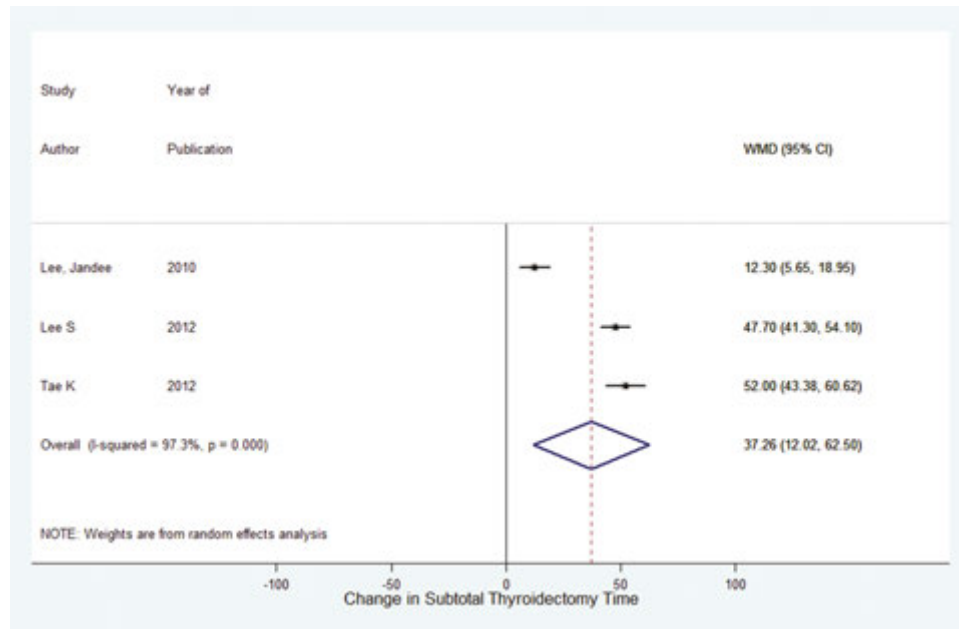


Figure 4. Subtotal thyroidectomy operative time, robotic vs open approach. CI, confidence interval; WMD, weighted mean difference.

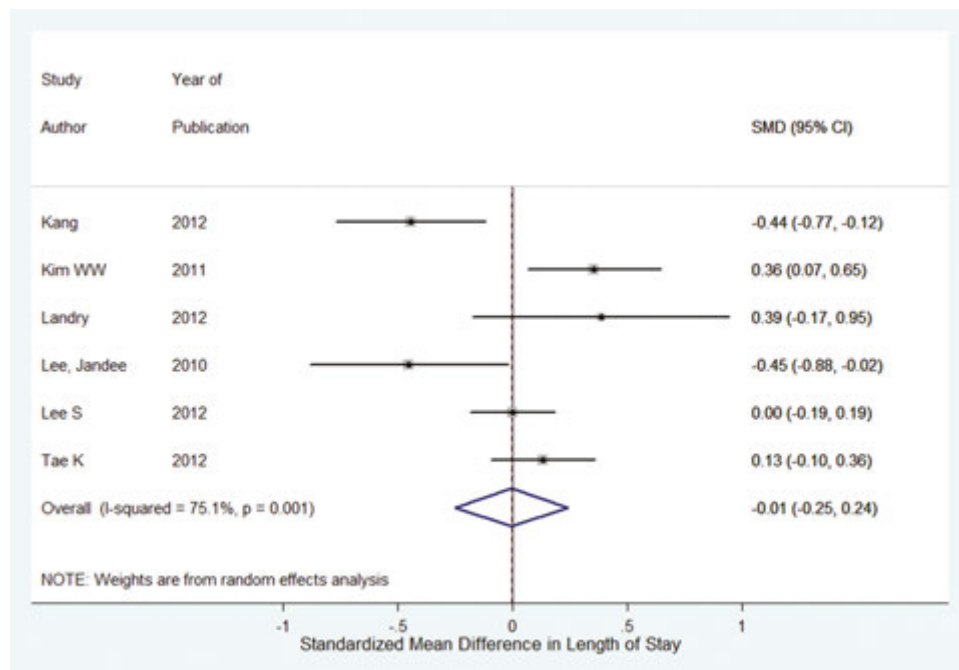


Figure 5. Hospital length of stay, robotic vs open approach. CI, confidence interval; SMD, standardized mean difference.

and handwriting due to brachial plexus injury; both patients recovered within several months. Lee et al²³ reported brachial plexus injuries in 5 of 41 (12.2%) patients, as manifested by shoulder discomfort, 1 week postoperatively; symptoms resolved in all patients by 3 months.

Quality of Life

Voice. Lee et al²³ used the Voice Handicap Index²⁸ preoperatively and 1 week and 3 months postoperatively for patient

self-assessment of voice impairment. No statistically significant difference in voice symptom scores was found at any time point between the 2 groups.

In a separate study, Lee et al²⁵ used both subjective and objective methods to evaluate patients preoperatively and 1 week and 3 months postoperatively. Subjective evaluation, performed by an experienced voice therapist using the GRBAS (Grade, Roughness, Breathiness, Asthenia, Strain) scale,²⁹ found no difference at any of the 3 time points. Objective evaluation was performed using videolaryngostroboscopy

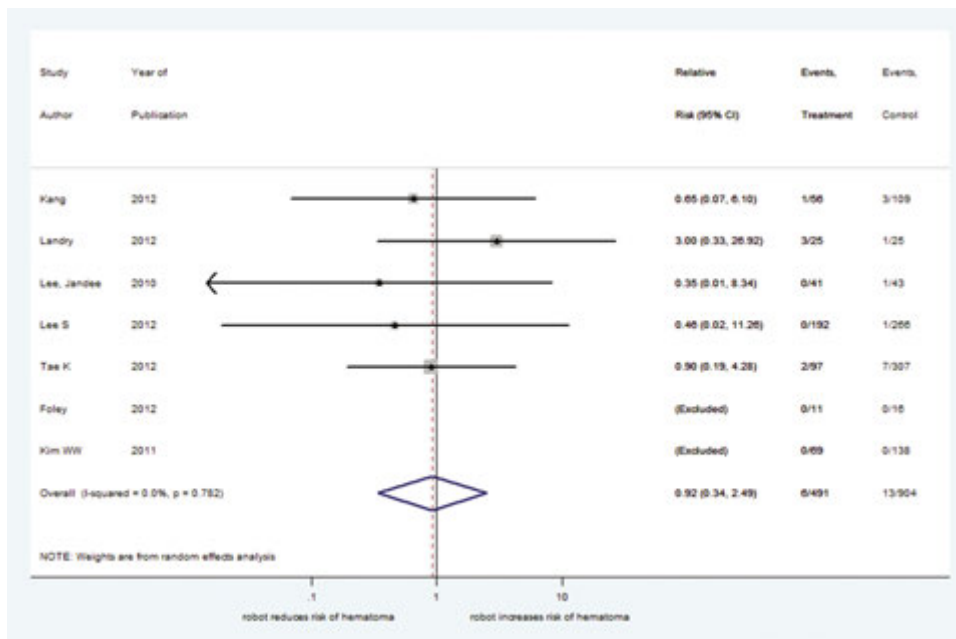


Figure 6. Risk of hematoma, robotic vs open approach. CI, confidence interval.

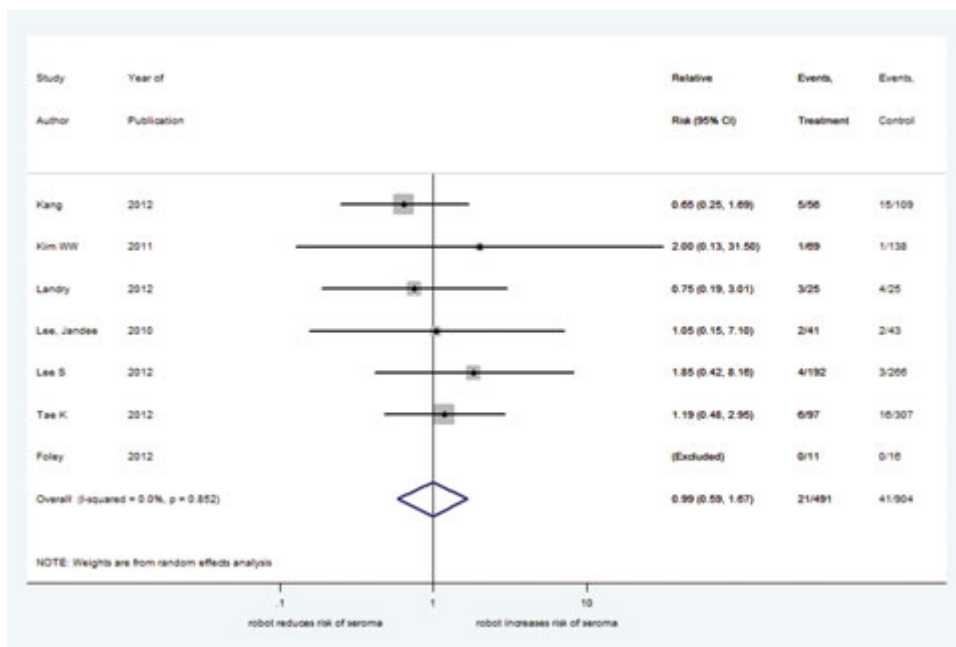


Figure 7. Risk of seroma, robotic vs open approach. CI, confidence interval.

(VLS), acoustic voice analysis, the Voice Range Profile (maximum and minimum frequencies, frequency range, intensity range, and maximum phonation time), and electroglottography (EGG). There were very few between-group differences identified, and all were temporary. In the acoustic voice analysis, the only significant difference was increased shimmer in the robotic group at 1 week, which resolved by 3 months. In addition, the aerodynamic studies and EGG revealed no notable differences at any time.

Tae et al¹⁸ used a 5-item questionnaire to investigate symptoms of vocal fatigue, hoarseness, pitch limitation, breathiness or weakness, and difficulty singing, with patients rating the severity of each symptom on a 5-point scale (none, mild, moderate, severe, or very severe). Preoperatively, the open cohort and the robotic cohort were similar in their reporting of vocal impairment. Postoperatively, the open cohort had a higher severity of voice symptoms at postoperative day 1 ($P = .008$), month 1 ($P = .049$), and month 3 ($P = .043$) but not at month 6.

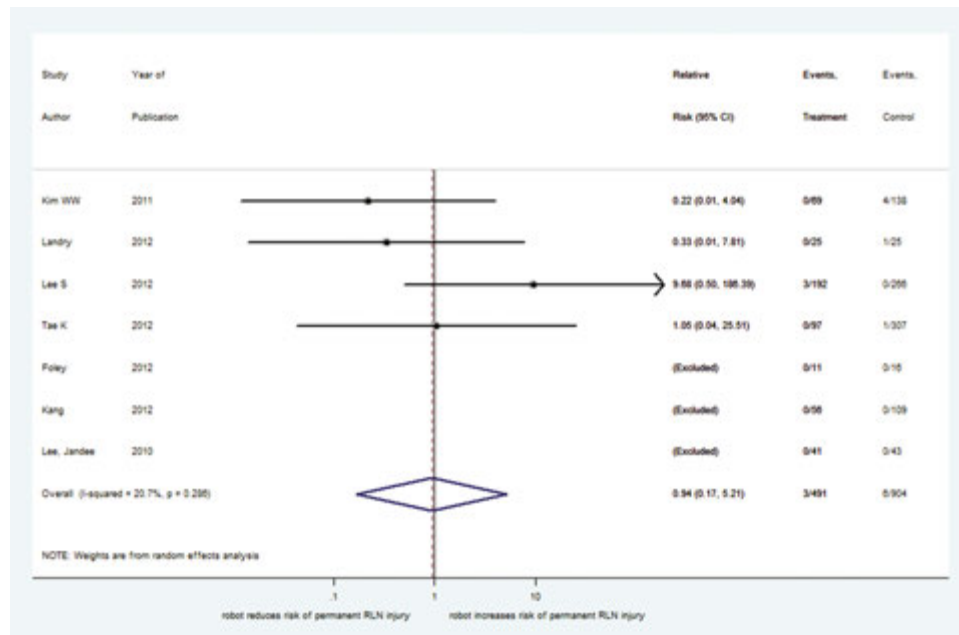


Figure 8. Risk of permanent recurrent laryngeal nerve (RLN) injury, robotic vs open approach. CI, confidence interval.

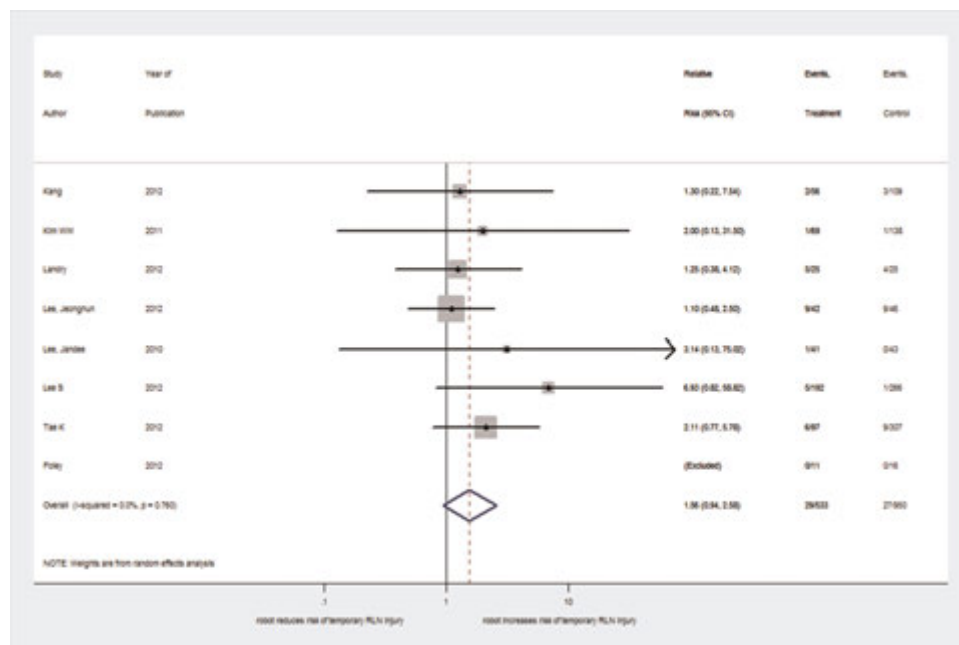


Figure 9. Risk of temporary recurrent laryngeal nerve (RLN) injury, robotic vs open approach. CI, confidence interval.

Objective studies included VLS, acoustic analysis, and the Voice Range Profile. The VLS results were not reported. Acoustic analysis revealed no significant differences between the 2 groups at any time point. Voice Range Profile parameters found a temporary broader frequency range and a higher maximum frequency at 3 months in the robotic cohort. This difference resolved by 6 months postoperatively. All 111 patients completed the questionnaire and underwent acoustic voice analysis preoperatively and at 1 week postoperatively. However, due to loss to follow-up,

only 105 patients completed the questionnaire at 1 month, 97 at 3 months, and 92 at 6 months, while 102 patients completed voice analysis at 1 month, 92 at 3 months, and 85 at 6 months postoperatively.

Swallowing. Swallowing was described in 2 articles with discordant results. Lee et al²³ used the validated 6-item Swallowing Impairment Index (SIS-6) for the assessment of increased effort of swallowing, choking, and throat clearing. Patients in both the robotic and open cohorts had similar

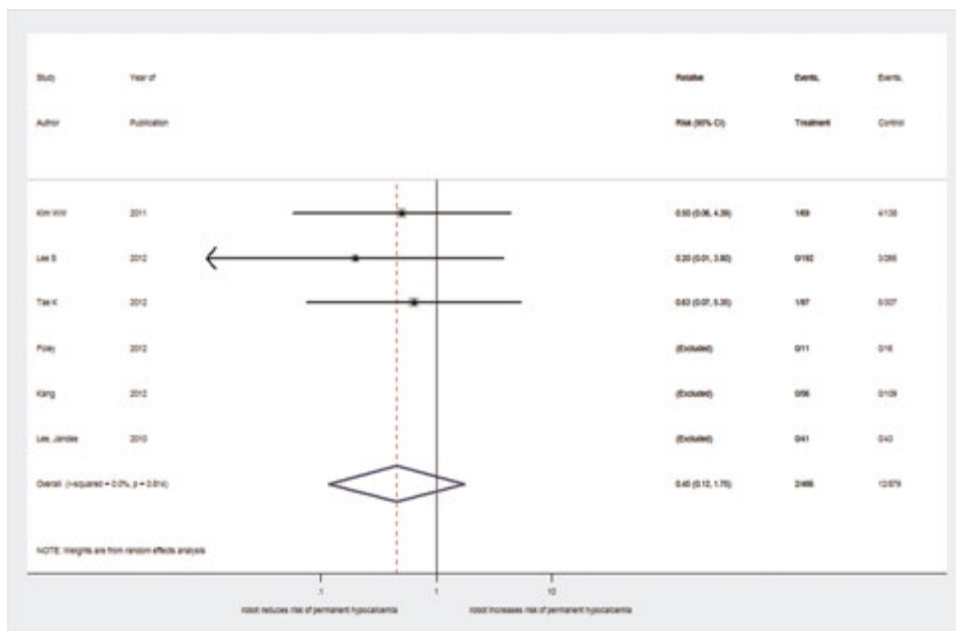


Figure 10. Risk of permanent hypocalcemia, robotic vs open approach. CI, confidence interval.

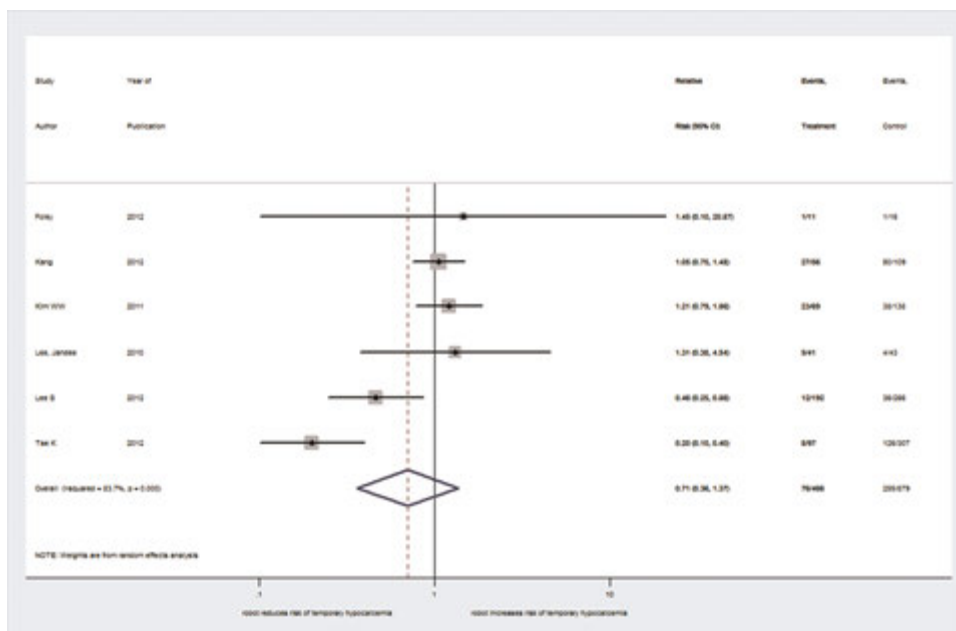


Figure 11. Risk of temporary hypocalcemia, robotic vs open approach. CI, confidence interval.

scores preoperatively. Postoperatively, the open cohort had significantly worse mean SIS-6 scores at both 1 week ($P = .001$) and 3 months ($P = .007$).

Tae et al¹⁸ created a 3-item questionnaire to evaluate pain or difficulty swallowing, foreign body sensation, and choking or coughing when swallowing. Patients reported their severity of symptoms on a 5-point scale (no symptoms, mild, moderate, severe, or very severe). There were no between-group differences in swallowing impairment at any time point measured up to 6 months.

Cosmetic satisfaction. Three articles described cosmetic satisfaction using nonvalidated instruments. Lee et al²³ assessed cosmetic satisfaction at 3 months postoperatively using a 5-point scale (extremely satisfied, satisfied, acceptable, dissatisfied, or extremely dissatisfied) and found that patients in the robotic group reported significantly greater satisfaction than those in the open group ($P < .0001$). Twenty-four (58.5%) patients in the robotic group reported being extremely satisfied compared with 5 (11.6%) patients in the open group. No patients in the robotic group reported being

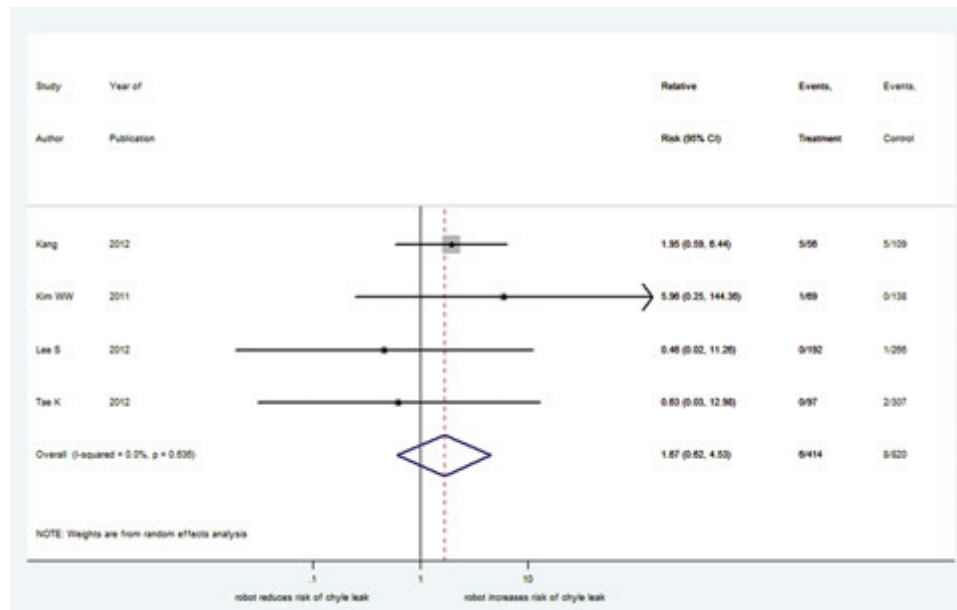


Figure 12. Risk of chyle leak, robotic vs open approach. CI, confidence interval.

either extremely dissatisfied or dissatisfied, while 8 patients (18.6%) in the open group reported being dissatisfied and 1 (2.3%) reported being extremely dissatisfied.

Tae et al¹⁷ assessed cosmetic satisfaction using a 5-point scale (very satisfied to very dissatisfied) at 1 week, 1 month, and 3 months postoperatively. Mean satisfaction was significantly better in the robotic group ($P < .001$ for each time point). However, the 2 study arms suffered from varying degrees of loss to follow-up; 32 of 41 (78.0%) patients in the robotic group reported cosmetic satisfaction, compared with only 89 of 163 (54.6%) in the open thyroidectomy group.

In 2012, Tae et al¹⁶ measured cosmetic satisfaction using the same 5-point scale at 1 day, 1 week, 1 month, and 3 months postoperatively. Mean satisfaction was significantly better in the robotic group than in the open group ($P < .001$ for each time point), but this study also suffered from differential loss to follow-up between the 2 study arms. In the robotic group, 48 of 75 (64.0%) patients reported cosmetic satisfaction scores, compared with 71 of 226 (31.4%) in the open thyroidectomy group.

Pain. Three articles described postoperative pain. Lee et al²³ used a 5-point grading scale to grade neck and anterior chest pain at 24 hours postoperatively. Pain was then reported as none, very slight, slight, moderate, or severe. Moderate or severe pain was reported by 5 (12.2%) patients in the robotic group and 6 (14.0%) in the open group ($P = .43$). The authors used an identical postoperative analgesic regimen for the 2 groups and found no difference in the number of postoperative analgesics used in each cohort ($P = .76$).

In the 2011 study by Tae et al,¹⁷ 121 patients scored their postoperative neck and anterior chest pain on a 5-point scale (none, mild, moderate, severe, or very severe) at 1 week, 1 month, and 3 months postoperatively. There

was no significant difference in the severity of neck pain. However, the robotic cohort reported significantly higher anterior chest pain scores at 1 week ($P < .001$). This difference resolved by 1 month postoperatively. In 2012, the same authors¹⁶ reported 119 patients who scored postoperative neck and anterior chest pain using the same 5-point scale but at 4 time points: 1 day, 1 week, 1 month, and 3 months postoperatively. There was again no significant difference in neck pain at any time point measured, but the robotic group had significantly higher anterior chest pain scores at 1 day, 1 week, and 1 month postoperatively ($P < .001$, $P < .001$, and $P = .014$, respectively). This difference in chest pain resolved by the 3-month postoperative visit.

Paresthesia. Three articles reported postoperative paresthesia in patients. Foley et al¹⁹ reported symptoms in 2 of 11 (18.2%) patients who underwent robotic thyroidectomy, both of whom recovered by the first postoperative visit. Lee et al²³ assessed for the presence of paresthesia or hyperesthesia of the neck with a questionnaire completed by all participants at 1 week and 3 months postoperatively. Significantly more patients reported neck paresthesia in the open cohort than in the robotic cohort, both at 1 week (95.3% vs 36.6%, $P = .010$) and at 3 months (65.1% vs 9.8%, $P = .002$) postoperatively. Anterior chest paresthesia at 1 week was reported by 19 (46.3%) patients in the robotic group and 1 (2.4%) in the open group. At 3 months, 8 (19.5%) patients in the robotic group and none in the open group were still experiencing chest paresthesia.

In the 2012 study by Tae et al,¹⁶ 119 patients reported paresthesias using a 4-point scale at 1 day, 1 week, 1 month, and 3 months postoperatively. The severity of neck paresthesias was significantly different at 3 months only, with the open group reporting a greater degree of neck

paresthesia ($P = .012$). The severity of anterior chest paresthesias was significantly greater in the robotic group 1 day, 1 week, and 1 month after surgery ($P = .001$, $P < .001$, and $P = .006$, respectively). This difference resolved by 3 months postoperatively.

Discussion

The current meta-analysis supports many of the findings of the previous meta-analysis by Jackson et al.⁸ We found no difference in the rates of RLN injury, hypocalcemia, hematoma, seroma, and chyle leakage. We also found that robotic thyroidectomy cases had significantly longer operative times (76.7 minutes in our study vs 42.05 minutes in Jackson et al⁸). To test whether this operative time discrepancy was due to robotic operative experience, we dropped the 2 US studies^{19,22} since South Korean surgeons generally have more experience with the surgical robot. This did not change the difference in operative times substantially (72.1 minutes longer in robotic thyroidectomy cases). Further analysis of the impact of the learning curve and institutional experience with the surgical robot on patient outcomes is warranted.³⁰

We caution that none of the studies in the meta-analysis were designed as a priori noninferiority or equivalence trials. Consequently, noninferiority of the experimental robotic technique cannot necessarily be assumed.^{31,32} This is supported in part by the fact that although many of the measured surgical outcomes were similar, the robotic approach has additional operative risks that have no analogue in traditional open surgery. Such unique complications can also have additional downstream adverse effects on QoL, such as brachial plexus injury and its impact on the ability to perform manual labor. Future research examining the impact of this particular complication on overall QoL would be informative.

Other discordant results from our meta-analysis highlight the complexity of evaluating the robotic thyroidectomy technique. For example, our study did not detect a significant difference in the overall LOS between the robotic and open cohorts, in contrast to Jackson et al.⁸ Since the previous meta-analysis included only studies conducted in Asia and the average LOS for all causes in South Korea is approximately triple that of the United States,³³ we conducted a sensitivity analysis excluding the US study by Landry et al.²² We determined that the SMD between open and robotic thyroidectomy hospital LOS was -0.053 (95% CI, -0.32 to 0.21), still a nonsignificant difference in mean LOS. This persistent difference might be due to the previous meta-analysis having double-counted overlapping cohorts by Tae et al,^{16,17} a problem we circumvented by acquiring the entire population of patients directly from the author and pooling them together into a single group.

One key issue is the impact of robotic thyroidectomy on cosmetic outcomes. Conceptually, robotic thyroidectomy might be expected to produce better cosmetic outcomes due to placement of the incision in a less visible area, and patients most likely to undergo thyroidectomy are younger

women who may place a higher premium on cosmesis. Indeed, the cohort undergoing robotic thyroidectomy in this meta-analysis were nearly all female and almost 9 years younger on average than the cohort undergoing open thyroidectomy. Moreover, the studies we identified suggested that cosmetic satisfaction was significantly higher in the patients undergoing robotic thyroidectomy.

However, these cosmetic results should be considered carefully in light of several methodological concerns. None of the studies on cosmesis were randomized or used validated instruments such as the Patient Scar Assessment Questionnaire,³⁴ which would standardize patient responses and allow more appropriate, less biased evaluation of the effectiveness of robotic vs open approaches. Since none of the studies were randomized, all are susceptible to self-selection bias; in other words, patients selecting a robotic treatment approach may be inherently biased toward greater cosmetic satisfaction because they selected that approach specifically to avoid a neck incision. Moreover, the results of the studies by Tae et al^{16,17} were confounded by attrition bias. Another important consideration is that none of the studies on cosmesis followed patients beyond 3 months. In typical scar formation, at the 3-month point, scar width is generally wider than the original incision and erythema is still present beyond the boundaries of the incision.³⁵ Since scar maturation tends to occur over a period of 1 year,³⁶ longer follow-up periods in future studies would be needed to determine whether the immediate impact of improved cosmesis using the surgical robot persists.

Our results should be viewed in the context of several potential limitations. First, the cohorts being compared in this meta-analysis were significantly different in many aspects: the patients in the robotic group were younger, had a lower BMI, and proportionally had fewer total thyroidectomies done in comparison to the open group. There are possible differences between American and South Korean patients in terms of body habitus, personal and physician preferences, and so on that may have affected patient selection within the studies we examined in our systematic review and meta-analysis. However, as a whole, the patients in the robotic cohort appear to be lower-risk surgical patients compared with the open cohort.

Methodologically, the current study is at risk of being affected by publication bias, since it is composed primarily of small observational studies demonstrating positive results, and studies with positive results are more likely to be published than those with negative results.³⁷⁻³⁹ Statistical heterogeneity is another concern; 5 of the 11 quantitative comparisons we performed had $I^2 > 75\%$, which is suggestive of high heterogeneity. This likely reflects the relatively few studies included in our meta-analysis, as well as fundamental differences in the populations being studied within each publication (eg, unilateral vs bilateral thyroid surgery, transaxillary vs combined transaxillary-breast approaches). On the other hand, our largest individual meta-analysis on patient outcomes included only 7 individual studies

(temporary RLN injury), and such small numbers of studies may be too low powered to reflect true heterogeneity.⁴⁰ Thus, caution is advised when evaluating the impact of high heterogeneity on the results of the meta-analysis.

Last, a valid cost analysis remains a key gap in the current literature. We identified 2 studies^{13,14} describing cost but were excluded from analysis because of methodological flaws. Broome et al¹³ showed a 217% increase in cost using robotic techniques, but the robotic thyroidectomy cost data were based on published series done at other institutions. Consequently, the cost of robotic surgery may not be reflective of the study institution itself. Cabot et al¹⁴ found significantly higher costs in robotic thyroidectomy compared with other approaches but only evaluated cost for the South Korean data in the study. Because most studies in this meta-analysis were conducted in South Korea, where reimbursement for robotic thyroidectomy is quadruple that for open thyroidectomy,⁴¹ the implications for US-based reimbursement strategies are uncertain. Robotic surgery also generates higher costs from related equipment usage, prolonged operative time, and associated facility and staffing fees. Thus, in the context of escalating health care costs in the United States, this represents a major concern at the national policy level.⁴²

Conclusion

Transaxillary and axillo-breast robotic thyroidectomy approaches have complication rates similar to the conventional transcervical approach. However, these robotic approaches also introduce unique complications that should not be ignored. Robotic thyroidectomy has been reported to improve cosmetic outcomes, although long-term studies using validated instruments will be essential to further demonstrate the value of the robotic approach. Valid, well-designed studies on the comparative costs of robotic and open thyroidectomy remain a critical gap in the research literature.

Acknowledgments

The authors thank Christine Landry, MD, Nancy Perrier, MD, and Kyung Tae, MD, for providing supplemental institutional data on robotic and open thyroidectomy patient outcomes and Ron Koppersmith, MD, for guidance on study design and critical review of our manuscript.

Author Contributions

Gordon H. Sun, substantial contributions to conception and design, acquisition of data, analysis and interpretation of data, drafting the article and revising it critically for important intellectual content, final approval of the version to be published; **Lilia Peress**, substantial contributions to conception and design, acquisition of data, analysis and interpretation of data, drafting the article and revising it critically for important intellectual content, final approval of the version to be published; **Melissa A. Pynnonen**, substantial contributions to conception and design, analysis and interpretation of data, drafting the article and revising it critically for important intellectual content, final approval of the version to be published.

Disclosures

Competing interests: The Robert Wood Johnson Foundation and the Department of Veterans Affairs were not directly involved in study design, data acquisition and interpretation, or manuscript preparation or review. Any opinions expressed herein do not necessarily reflect the opinions of the Robert Wood Johnson Foundation or the Department of Veterans Affairs. At the time of manuscript preparation, Gordon H. Sun was a Robert Wood Johnson Foundation Clinical Scholar supported by the US Department of Veterans Affairs. Dr Sun is now an employee of Partnership for Health Analytic Research, LLC in Beverly Hills, California, and a general otolaryngologist at the UCLA Arthur Ashe Student Health & Wellness Center in Los Angeles, California.

Sponsorships: Robert Wood Johnson Foundation Clinical Scholars Program, US Department of Veterans Affairs.

Funding source: None.

Supplemental Material

Additional supporting information may be found at <http://oto.sagepub.com/content/by/supplemental-data>

References

- Giddings AE. The history of thyroidectomy. *J R Soc Med.* 1998;91(suppl 33):3-6.
- Hegner CF. A history of thyroid surgery. *Ann Surg.* 1932;95:481-492.
- Huscher CS, Chiodini S, Napolitano C, Recher A. Endoscopic right thyroid lobectomy. *Surg Endosc.* 1997;11:877.
- Goh HK, Ng YH, Teo DT. Minimally invasive surgery for head and neck cancer. *Lancet Oncol.* 2010;11:281-286.
- Patel D, Kebebew E. Pros and cons of robotic transaxillary thyroidectomy. *Thyroid.* 2012;22:984-985.
- Koppersmith RB, Salem A, Holsinger FC. Advanced approaches for thyroid surgery. *Otolaryngol Head Neck Surg.* 2009;141:340-342.
- Lee J, Kang SW, Jung JJ, et al. Multicenter study of robotic thyroidectomy: short-term postoperative outcomes and surgeon ergonomic considerations. *Ann Surg Oncol.* 2011;18:2538-2547.
- Jackson NR, Yao L, Tufano RP, Kandil EH. Safety of robotic thyroidectomy approaches: meta-analysis and systematic review. *Head Neck.* 2014;36:137-143.
- Lin S, Chen ZH, Jiang HG, Yu JR. Robotic thyroidectomy versus endoscopic thyroidectomy: a meta-analysis. *World J Surg Oncol.* 2012;10:239.
- Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med.* 2009;151:264-269, W264.
- Sun GH, Peress L, Pynnonen MA. Meta-analysis of robotic vs. conventional thyroidectomy approaches for thyroid disease. National Institute for Health Research. 2013. www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42012003331. Accessed December 31, 2013.
- Martin RC II, Brennan MF, Jaques DP. Quality of complication reporting in the surgical literature. *Ann Surg.* 2002;235:803-813.

13. Broome JT, Pomeroy S, Solorzano CC. Expense of robotic thyroidectomy: a cost analysis at a single institution. *Arch Surg*. 2012;147:1102-1106.
14. Cabot JC, Lee CR, Brunaud L, et al. Robotic and endoscopic transaxillary thyroidectomies may be cost prohibitive when compared to standard cervical thyroidectomy: a cost analysis. *Surgery*. 2012;152:1016-1024.
15. 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;27:1321-1325.
16. Tae K, Ji YB, Cho SH, Lee SH, Kim DS, Kim TW. Early surgical outcomes of robotic thyroidectomy by a gasless unilateral axillo-breast or axillary approach for papillary thyroid carcinoma: 2 years' experience. *Head Neck*. 2012;34:617-625.
17. 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;25:221-228.
18. Tae K, Kim KY, Yun BR, et al. Functional voice and swallowing outcomes after robotic thyroidectomy by a gasless unilateral axillo-breast approach: comparison with open thyroidectomy. *Surg Endosc*. 2012;26:1871-1877.
19. Foley CS, Agcaoglu O, Siperstein AE, Berber E. Robotic transaxillary endocrine surgery: a comparison with conventional open technique. *Surg Endosc*. 2012;26:2259-2266.
20. Kang SW, Lee SH, Park JH, et al. A comparative study of the surgical outcomes of robotic and conventional open modified radical neck dissection for papillary thyroid carcinoma with lateral neck node metastasis. *Surg Endosc*. 2012;26:3251-3257.
21. Kim WW, Kim JS, Hur SM, et al. Is robotic surgery superior to endoscopic and open surgeries in thyroid cancer? *World J Surg*. 2011;35:779-784.
22. Landry CS, Grubbs EG, Warneke CL, et al. Robot-assisted transaxillary thyroid surgery in the United States: is it comparable to open thyroid lobectomy? *Ann Surg Oncol*. 2012;19:1269-1274.
23. 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;24:3186-3194.
24. Lee S, Ryu HR, Park JH, et al. Early surgical outcomes comparison between robotic and conventional open thyroid surgery for papillary thyroid microcarcinoma. *Surgery*. 2012;151:724-730.
25. Lee J, Na KY, Kim RM, et al. Postoperative functional voice changes after conventional open or robotic thyroidectomy: a prospective trial. *Ann Surg Oncol*. 2012;19:2963-2970.
26. 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;27:1587-1593.
27. Lee KE, Koo do H, Im HJ, et al. Surgical completeness of bilateral axillo-breast approach robotic thyroidectomy: comparison with conventional open thyroidectomy after propensity score matching. *Surgery*. 2011;150(6):1266-1274.
28. Rosen CA, Lee AS, Osborne J, Zullo T, Murry T. Development and validation of the voice handicap index-10. *Laryngoscope*. 2004;114:1549-1556.
29. Hirano M. Clinical examination of voice. In: Arnold GE, Winckel F, Wyke BD, eds. *Disorders of Human Communication*. New York, NY: Springer; 1981.
30. Gross ND. Commentary: is robotic thyroid surgery worth the learning curve? *Otolaryngol Head Neck Surg*. 2012;147:1047-1048.
31. Piaggio G, Elbourne DR, Pocock SJ, Evans SJ, Altman DG. Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. *JAMA*. 2012;308:2594-2604.
32. Treadwell JR, Uhl S, Tipton K, et al. Assessing equivalence and noninferiority. *J Clin Epidemiol*. 2012;65:1144-1149.
33. Organisation for Economic Co-operation and Development. *Health at a Glance 2011: OECD Indicators*. Paris, France: OECD Publishing; 2011.
34. Durani P, McGrouther DA, Ferguson MW. The Patient Scar Assessment Questionnaire: a reliable and valid patient-reported outcomes measure for linear scars. *Plast Reconstr Surg*. 2009;123:1481-1489.
35. Bond JS, Duncan JA, Sattar A, et al. Maturation of the human scar: an observational study. *Plast Reconstr Surg*. 2008;121:1650-1658.
36. Kokoska MS, Thomas JR. Scar revision. In: Papel ID, Frodel JL, Holt GR, et al, eds. *Facial Plastic and Reconstructive Surgery*. 3rd ed. New York, NY: Thieme; 2009:59-65.
37. Dwan K, Gamble C, Williamson PR, Kirkham JJ. Systematic review of the empirical evidence of study publication bias and outcome reporting bias—an updated review. *PLoS One*. 2013;8:e66844.
38. Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. *Lancet*. 1991;337:867-872.
39. Sterne JAC, Harbord RM. Funnel plots in meta-analysis. *Stata J*. 2004;4:127-141.
40. Ioannidis JP, Patsopoulos NA, Evangelou E. Uncertainty in heterogeneity estimates in meta-analyses. *BMJ*. 2007;335:914-916.
41. Dionigi G. Robotic thyroidectomy: Seoul is not Varese. *Otolaryngol Head Neck Surg*. 2013;148:178.
42. Barbash GI, Glied SA. New technology and health care costs—the case of robot-assisted surgery. *N Engl J Med*. 2010;363:701-704.