Premiere Publications from The Triological Society

Read all three of our prestigious publications, each offering high-quality content to keep you informed with the latest developments in the field.

**Laryngoscope**

*Founded in 1896*

Editor-in-Chief: Michael G. Stewart, MD, MPH

The leading source for information in head and neck disorders.

[**Laryngoscope.com**](https://www.laryngoscope.com)

**Investigative Otolaryngology**

Editor-in-Chief: D. Bradley Welling, MD, PhD, FACS

Rapid dissemination of the science and practice of otolaryngology-head and neck surgery.

[**InvestigativeOto.com**](https://www.investigativeoto.com)

**ENTtoday**

A publication of the Triological Society

Editor-in-Chief: Alexander Chiu, MD

Must-have timely information that Otolaryngologist-head and neck surgeons can use in daily practice.

[**Enttoday.org**](https://www.enttoday.org)
Systematic Review

Robotic Versus Conventional Neck Dissection: A Systematic Review and Meta-Analysis

Daniel C. Sukato, MD; Daniel P. Ballard, MD; Jason M. Abramowitz, MD; Richard M. Rosenfeld, MD, MPH; Stefan Mlot, MD

Objective: The aim of this systematic review is to compare the perioperative outcomes of robotic versus conventional neck dissection in patients with head and neck malignancy.

Methods: An electronic search of PubMed, Web of Science, and EMBASE databases was conducted. We included studies with direct comparisons of robotic and open neck dissections and performed dual, independent data extraction for primary outcomes of nodal yield, recurrence rate, subjective cosmetic assessment, operative time, length of stay, and rates of perioperative complications. Data were pooled using random effects meta-analysis to determine the standardized mean difference (SMD), absolute risk difference (RD), and 95% confidence intervals (CI).

Results: Eleven comparative studies comprising 225 robotic and 430 open neck dissections met the final selection criteria. All studies had low to moderate risk of bias. Robotic surgery improved cosmesis (SMD 1.15, 95% CI 0.73 to 1.56) but also increased operative time (SMD 1.94, 95% CI 1.25 to 2.63). Total nodal yield, pathological nodal yield, recurrence rate, rates of perioperative complications, and length of stay were not significantly different between the two groups, and the 95% CIs suggested that false negative results were unlikely. The results remained consistent after stratification by pathology and robotic technique.

Conclusion: Although robotic neck dissection may offer similar perioperative outcomes compared to conventional neck dissection, it requires significantly more operative time. Whereas cosmesis was found to be superior among the robotic cohort, this must be viewed cautiously given the nonvalidated measurement tool that was used and the inherent reporting bias associated with it.

Key Words: Robotic, neck dissection, cervical lymphadenectomy, squamous cell carcinoma, papillary thyroid carcinoma.

Laryngoscope, 129:1587–1596, 2019

INTRODUCTION

Robotic-assisted surgical procedures have been performed in a number of surgical subspecialties, often with comparable or superior outcomes relative to conventional approaches.1-6 The advantages of robotic surgery, which include improved dexterity, high-definition three-dimensional visualization, and the elimination of physiological tremors, give this approach the potential to overcome the limitations of human hands and laparoscopic instruments.7 Because acceptance of this technology has grown in the field of head and neck surgery, so too have its indications. Robotic approaches have been widely adopted in transoral surgery, most often utilized in procedures such as radical tonsillectomy, base-of-tongue resection, supraglottic laryngectomy, and inset of free tissue transfer, among others.8-21 The increasing demand for scarless neck surgery has prompted interest in the utilization of this technology in cervical lymphadenectomy. Kang et al.22,23 approached the neck robotically through a transaxillary incision, achieving adequate dissections of levels IIA, III, IV, and VB. Terris et al.24,25 popularized the facelift incision for robotic thyroidectomy, demonstrating improved operative time compared to the transaxillary technique. Following this, Kim et al.26 combined the facelift and transaxillary approach to perform neck dissections in head and neck squamous cell carcinoma. Multiple case series and comparative studies, primarily out of South Korea, subsequently highlighted the feasibility of nodal dissection through robotic transaxillary, retroauricular, and modified facelift approaches.22,27-40 Despite the embrace of this technology in East Asia, studies remain limited in the head and neck surgical literature. Given the increasing adoption and popularity of robotic surgery, it is imperative to systematically assess the evidence supporting the application of robotic techniques to cervical lymphadenectomy.
The objective of this systematic review is to compare the surgical outcomes of robotic and conventional neck dissection in patients with head and neck malignancy. We used meta-analysis to quantitatively pool and synthesize data from comparative studies. To our knowledge, this is the first systematic review to compare robotic and conventional neck dissections.

METHODS

The research question was framed in the population, intervention, comparison, and outcomes format. In adults who underwent neck dissection for head and neck malignancy, how do the perioperative outcomes of the robotic approach compare to those of the conventional approach? This systematic review was reported in adherence to Preferred Reporting Items for Systematic Reviews and Meta-Analyses standards.

An electronic query of PubMed, Web of Science, and EMBASE databases was conducted using the keywords (“robot” or “robotic”) and “neck dissection.” The inquiry was limited to English-only articles. Conference oral or poster abstracts, book chapters, and textbooks were removed. Studies that discussed other types of minimally invasive approaches or involved minimally invasive procedures in other organ systems were excluded. Only comparative studies between robotic and conventional neck dissection were included. In recognition of the variety of robotic approaches to the lateral neck, no limitations were placed on the robotic device or how the robotic procedure was performed. For studies with overlapping patient cohorts, the study with the highest quality and largest sample size was selected for data analysis. Additionally, only outcomes that were investigated by four or more studies were evaluated.

Two independent investigators (D.C.S., J.M.A.) reviewed and screened abstracts and full-text articles by applying the aforementioned inclusion and exclusion criteria. Using a standardized extraction sheet, the two reviewers then independently extracted data from the included studies, specifically obtaining sample size, demographic variables, length of follow-up, oncologic parameters, cosmetic satisfaction score, operative time, length of hospital stay, and rates of intra- and postoperative complications. Assessments of bias and methodological quality of the studies were performed by utilizing the Methodological Index for Non-Randomized Studies (MINORS) criteria. Any discrepancies and disagreements encountered during the dual-investigator review were resolved through a discussion with the senior author of this study (S.M.).

Statistical analysis was conducted by utilizing meta-analysis software to pool the primary outcomes between robotic and conventional neck dissections. For continuous variables, pooled weighted means and standard deviations (SD) were estimated based on study sample sizes. Effect size was reported as the standardized mean difference (SMD), which was calculated as the mean difference in outcome between two groups divided by the pooled SD. The SMD is comparable across studies, with zero indicating no difference and higher numbers reflecting larger differences. Additionally, the SMD is similar to Cohen’s effect size, and its values can be considered trivial if < 0.20, small if 0.20 to 0.49, moderate if 0.50 to 0.79, and large if 0.80 or higher. For binomial outcomes, the absolute risk difference (RD) was reported and was calculated as the difference in incidence between the conventional and robotic group. If a study offered outcomes in medians and/or ranges, the means and SDs were estimated as described by Wan et al. All data were pooled using random effects meta-analysis, which assumes the existence of multiple true effect sizes across studies and results in wider 95% confidence intervals (CI) compared to the fixed effects approach. Heterogeneity was evaluated using I² statistics, where values of 25%, 50%, and 75% correspond to low, moderate, and high heterogeneity, respectively. All comparisons were assessed for significance with a type I error probability (P value) threshold of 0.05.

The analysis was further stratified based on pathology and robotic technique. All patients with papillary thyroid carcinoma (PTC) underwent the robotic approach with a form of transaxillary incision, whereas all patients with either squamous cell carcinoma or other non-PTC pathology underwent a form of retroauricular incision. Therefore, we classified the robotic approach into two major techniques: transaxillary and retroauricular. The gasless unilateral axillary approach, gasless unilateral axillo-breast approach, bilateral axillo-breast approach, and combined transaxillary and retroauricular approach were designated together as the transaxillary technique. The retroauricular, postauricular, and modified facelift approaches comprised the retroauricular technique.

RESULTS

The initial query yielded 1,496 studies, of which 1,245 studies remained after duplicates were excluded.
The remaining studies underwent a dual investigator screen, which resulted in 16 articles for full-text reviews and 11 studies for final analysis (Fig. 1). Characteristics of the studies, associated demographic variables, and initial tumor parameters are listed in Table I. Ten of the studies originated from South Korea, whereas one of the studies was from Pittsburgh, Pennsylvania. The most common type of study was prospective (n = 6), followed by retrospective or case series with chart reviews (n = 5). None of the studies had randomization performed between the two groups. Six of the studies involved patients with head and neck malignancy who underwent a retroauricular approach to the neck, whereas five of the studies consisted of patients with PTC who underwent a transaxillary approach. The total sample sizes ranged from nine to 143, whereas the robotic and conventional group sample sizes spanned from three to 62 and six to 102, respectively. The pooled mean age ranged from 36.7 to 61.4 for the robotic group and from 43.5 to 64 for the conventional group. The pooled body mass index (BMI) was obtained from five studies and ranged from 36.7 to 61.4 for the robotic group and from 43.5 to 62 and six to 102, respectively. The pooled mean lengths of follow-up timelines varied from 6 to 60 months. Three studies did not specify lengths of follow-up. In seven studies, patients underwent selective neck dissection (SNd). In four studies, patients were subjected to modified radical neck dissection. Three studies performed neck dissection electively on the N0 neck, whereas eight studies included only therapeutic neck dissections in their patient cohorts. All studies utilized the d Vinci Surgical System (Intuitive Surgical, Inc., Sunnyvale, CA) in their robotic-assisted interventions.

The methodological quality of each study was evaluated in accordance to MINORS criteria (Table II). The MINORS scores ranged from 18 to 21, with 24 as the perfect score for comparative studies. The greater the value of the MINORS score, the lower the risk of inherent bias. The mean and median MINORS scores were 19.6 and 17.5, respectively. The majority of studies were deficient in power calculations and blinded evaluation of outcomes. Lee et al. had the assessment of nodal yield blinded on pathology specimens but did not blind the evaluation of other outcomes. Two studies had significant group differences in certain baseline demographic characteristics, specifically age, BMI, and gender distribution. Although five studies were retrospective in nature, all involve prospectively collected data with an a priori protocol. Notably, in all comparative studies, patient financial status and preference strongly influenced selection for robotic procedures.

**Operative Time and Length of Stay**

Ten studies evaluated and compared operative times between robotic and conventional groups (Fig. 2A). The pooled mean estimates were 254.14 (SD 47.7) and 189.16 (SD 54.6) minutes, respectively. Random effects meta-

---

**TABLE I. Characteristics of Studies Included in the Systematic Review and Meta-Analysis**

<table>
<thead>
<tr>
<th>Study Year</th>
<th>Country</th>
<th>Study Type</th>
<th>Sample Size N</th>
<th>Age Mean (SD)</th>
<th>BMI Mean (SD)</th>
<th>T3 Stage %</th>
<th>T4 Stage %</th>
<th>Robotic Approach</th>
<th>Conventional Approach</th>
<th>Type of ND</th>
<th>Neck Stage</th>
<th>MINORS Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012 SK</td>
<td>SK P</td>
<td>21 10</td>
<td>16 44 (11.8)</td>
<td>56.5 (13.6)</td>
<td>N/S</td>
<td>0.0</td>
<td>6.3</td>
<td>SND</td>
<td>N/S</td>
<td>N0</td>
<td>21</td>
<td>24</td>
</tr>
<tr>
<td>2013 SK</td>
<td>SK P</td>
<td>18 7</td>
<td>24 58 (6.8)</td>
<td>N/S</td>
<td>26.6</td>
<td>0.0</td>
<td>10.5</td>
<td>PAFL MRND</td>
<td>MFL RA</td>
<td>N1b</td>
<td>20</td>
<td>22</td>
</tr>
<tr>
<td>2013 USA</td>
<td>USA R</td>
<td>19 11</td>
<td>32 53 (8.1)</td>
<td>N/S</td>
<td>16.7</td>
<td>0.0</td>
<td>10.5</td>
<td>PAFL MRND</td>
<td>MFL RA</td>
<td>N1b</td>
<td>20</td>
<td>22</td>
</tr>
<tr>
<td>2014 SK</td>
<td>SK R</td>
<td>20 11</td>
<td>25 60.75 (8)</td>
<td>N/S</td>
<td>10.5</td>
<td>0.0</td>
<td>10.5</td>
<td>PAFL MRND</td>
<td>MFL RA</td>
<td>N1b</td>
<td>20</td>
<td>22</td>
</tr>
<tr>
<td>2015 SK</td>
<td>SK P</td>
<td>20 13</td>
<td>41 64 (11.7)</td>
<td>N/S</td>
<td>10.5</td>
<td>0.0</td>
<td>10.5</td>
<td>PAFL MRND</td>
<td>MFL RA</td>
<td>N1b</td>
<td>20</td>
<td>22</td>
</tr>
<tr>
<td>2015 SK</td>
<td>SK R</td>
<td>20 11</td>
<td>25 62 55.4 (8.9)</td>
<td>10.5</td>
<td>0.0</td>
<td>10.5</td>
<td>PAFL MRND</td>
<td>MFL RA</td>
<td>N1b</td>
<td>20</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>2016 SK</td>
<td>SK P</td>
<td>20 11</td>
<td>25 62 55.4 (8.9)</td>
<td>10.5</td>
<td>0.0</td>
<td>10.5</td>
<td>PAFL MRND</td>
<td>MFL RA</td>
<td>N1b</td>
<td>20</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>2016 SK</td>
<td>SK R</td>
<td>20 11</td>
<td>25 62 55.4 (8.9)</td>
<td>10.5</td>
<td>0.0</td>
<td>10.5</td>
<td>PAFL MRND</td>
<td>MFL RA</td>
<td>N1b</td>
<td>20</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>2017 SK</td>
<td>SK P</td>
<td>20 11</td>
<td>25 62 55.4 (8.9)</td>
<td>10.5</td>
<td>0.0</td>
<td>10.5</td>
<td>PAFL MRND</td>
<td>MFL RA</td>
<td>N1b</td>
<td>20</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>2017 SK</td>
<td>SK R</td>
<td>20 11</td>
<td>25 62 55.4 (8.9)</td>
<td>10.5</td>
<td>0.0</td>
<td>10.5</td>
<td>PAFL MRND</td>
<td>MFL RA</td>
<td>N1b</td>
<td>20</td>
<td>22</td>
<td></td>
</tr>
</tbody>
</table>

BABA: bilateral axillo-breast approach; BMI: body mass index; GUA: gasless unilateral; GUAB: gasless unilateral axillo-breast approach; L: left; M: modified; RA: retroauricular; R: right; SA: squamous cell carcinoma; SD: standard deviation; SK: South Korea; SND: selective neck dissection; TARA: transaxillary retroauricular; USA: United States of America.
analysis demonstrated a statistically significant longer operative time for the robotic approach compared to the conventional approach, with a total SMD of 1.94 (95% CI 1.25 to 2.63). The 95% CI suggested high confidence with a large effect size. $I^2$ of 89% indicated a high degree of heterogeneity among the studies. When the transaxillary and retroauricular approaches were analyzed separately, the operative times remained significantly prolonged in the robotic group compared to the conventional study population. The effect size was larger for the retroauricular method (SMD of 2.63) compared to the transaxillary technique (SMD of 1.67).

Length of stay was evaluated in eight studies, and the associated forest plot is shown in Figure 2B. Although the SMD favored the robotic group with a trivial effect size, the CI was not able to exclude possible outcomes in favor of the conventional group (SMD of $-0.148$, 95% CI $-0.364$ to 0.068). The pooled mean length of stay estimates were 8.28 (SD 3.94) versus 8.74 (SD 4.67) days, respectively. When stratified by robotic techniques, the differences in length of stay remained nonsignificant, with similar effect sizes.

**Perioperative Results**

Perioperative parameters and complications including chyle leak, hematoma, Horner’s syndrome, marginal mandibular nerve weakness, seroma, wound infection, and surgical wound drainage were assessed by multiple comparative studies (Supporting Fig. S1A–G). No significant differences were noted between robotic and conventional groups; and all analyses had low heterogeneity. These findings, especially with regard to chyle leak, hematoma, Horner’s syndrome, marginal mandibular nerve weakness, seroma, and wound infection, were associated with narrow 95% CIs, suggesting sufficient statistical power to exclude false negative results. The SMD for surgical drainage (SMD 0.179, 95% CI of $-0.134$ to 0.492) favored the conventional group with a trivial effect size, but the CI was not able to exclude potential outcomes in favor of the robotic approach. The pooled mean total drainage amounts were 547.95 (SD 389.3) and 478.54 (SD 262.83) mL between robotic and conventional cohorts, respectively. Notably, the analyzed studies did not consistently provide duration of drainage or criteria for removal of drains.

**Total Lateral Nodal Yield, Pathological Yield, and Recurrence Rate**

Lateral nodal yield was defined as the amount of cervical lymph nodes obtained with the exclusion of central compartment nodes. Comparisons of total lateral nodal yield between robotic and conventional groups were performed in 10 studies. The pooled mean lateral nodal yields were 31.46 (SD 11.27) for the robotic cohort and 31.53 (SD 13.37) for the open group. A random effects meta-analysis is demonstrated in a forest plot (Fig. 3A). Standardized mean difference of 0.003 (95% CI of $-0.181$ to 0.186, $I^2 = 20\%$) was obtained, which indicated no statistically significant difference in lateral nodal yield between the two approaches. The narrow range of the CI demonstrated adequate statistical

<table>
<thead>
<tr>
<th>Study (Year)</th>
<th>Score*</th>
<th>Clearly Stated Aim</th>
<th>Inclusion of Consecutive Patients</th>
<th>Prospective Collection of Data</th>
<th>Appropriate Calculation of Study Size</th>
<th>Unbiased Assessment of Study Endpoint</th>
<th>Loss to Follow-Up</th>
<th>Follow-Up Period to Aim of Study</th>
<th>Unbiased Assessment of Study Endpoint</th>
<th>Appropriate Follow-Up Period to Aim of Study</th>
<th>Prospective Calculation of Study Size</th>
<th>Adequate Statistical Analyses</th>
<th>Adequate Control Group</th>
<th>Contemporary Groups</th>
<th>Adequate Statistical Analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lee 2012</td>
<td>21</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Park 2013</td>
<td>18</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Albergotti 2014</td>
<td>18</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Tan 2014</td>
<td>19</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Kim 2016</td>
<td>20</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Kim 2017</td>
<td>20</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Song 2016</td>
<td>20</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Kim 2017</td>
<td>20</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Song 2016</td>
<td>20</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

*MINORS score ranges from 0 to 24. 1 if reported but inadequate, and 2 if reported and adequate.

<table>
<thead>
<tr>
<th>Study (Year)</th>
<th>Score*</th>
<th>Clearly Stated Aim</th>
<th>Inclusion of Consecutive Patients</th>
<th>Prospective Collection of Data</th>
<th>Appropriate Calculation of Study Size</th>
<th>Unbiased Assessment of Study Endpoint</th>
<th>Loss to Follow-Up</th>
<th>Follow-Up Period to Aim of Study</th>
<th>Unbiased Assessment of Study Endpoint</th>
<th>Appropriate Follow-Up Period to Aim of Study</th>
<th>Prospective Calculation of Study Size</th>
<th>Adequate Statistical Analyses</th>
<th>Adequate Control Group</th>
<th>Contemporary Groups</th>
<th>Adequate Statistical Analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lee 2012</td>
<td>21</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Park 2013</td>
<td>18</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Albergotti 2014</td>
<td>18</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Tan 2014</td>
<td>19</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Kim 2016</td>
<td>20</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Kim 2017</td>
<td>20</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Song 2016</td>
<td>20</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Kim 2017</td>
<td>20</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Song 2016</td>
<td>20</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

*MINORS score ranges from 0 to 24. 1 if reported but inadequate, and 2 if reported and adequate.

MINORS = Methodological Index for Non-Randomized Studies.
power. These findings remained consistent after stratification based on robotic techniques.

Pathological lateral nodal yield was also compared between the two groups, and a random effects meta-analysis from six studies is demonstrated in Supporting Figure S2. No significant difference was noted (SMD $-0.038$, 95% CI of $-0.247$ to $0.171$, $I^2 = 0\%$). The pooled mean pathological nodal yield was 4.4 (4.13) and 4.53 (3.89) between the robotic and open groups, respectively.

Recurrence rates at follow-ups of at least 6 months were evaluated in seven studies. Random effects meta-analysis is shown Figure 3B. No statistically significant difference in recurrence rates (RD of $-0.004$, 95% CI of $-0.028$ to $0.019$, $I^2 = 4\%$) was found, and the tight CI suggested sufficient statistical power. After stratification of the analysis by robotic techniques, there continued to be no significant difference in 6-month recurrence rates between robotic and open approaches to the lateral neck.

**Cosmesis**

Cosmesis was evaluated through a nonvalidated 5-point subjective scoring tool in five studies. These
evaluations were performed at least 3 months postoperatively, with a score of 1 or 5 considered as “extremely satisfied” depending on the study. The scoring systems were converted accordingly, with a score of 5 corresponding to highest satisfaction. A random effects meta-analysis was performed and is demonstrated in a forest plot (Fig. 4). Lower cosmetic scores were associated with conventional surgeries, supported by a large effect size (SMD −1.15, 95% CI −1.56 to −0.73). The narrow 95% CI was consistent with a high statistical power. However, an I² statistic of 67% indicated moderate to high heterogeneity. When the analysis was stratified based on robotic approach, the transaxillary method was associated with superior cosmesis compared to open procedures (SMD of −1.12, 95% CI −1.55 to −0.69). This did not reach significance in the retroauricular cohort (P = 0.068).

**DISCUSSION**

The conventional, or open, neck dissection is the gold standard surgical approach for elective or therapeutic lymphadenectomy in patients with head and neck cancer. However, it is important to acknowledge that the goals of any particular surgical intervention are distinct from the specific techniques employed to achieve those goals. New surgical approaches may offer unique advantages compared to conventional approaches, although these benefits must be considered in concert with their limitations. It is
important to investigate if the novel approach offers equivalent or improved outcomes, especially with regard to rates of potential complications, expenditure of healthcare resources, postoperative quality of life, and the ability to achieve the goals of surgical care. In this systematic review, we found that, compared to conventional neck dissection, robotic neck dissection offers similar perioperative results but places a significantly higher burden on operative time. Due to the lack of validated assessments of cosmesis, no conclusion can be made at this time as to whether cosmesis is significantly different when comparing the two approaches.

Given that long-term survival data were lacking in the studies that comprised this analysis, true oncological outcomes were deficient in our study. This was perhaps due to the novelty of robotic neck dissection, which became clear when considering that the included studies were published in 2012 or after. Additionally, information on other prognostic factors, such as extranodal extension and rate of adjuvant chemotherapy and/or radiation, was not consistently provided. However, given the limited follow-up time frame of the studies, it was reasonable to use total lateral and pathological lymph node yield as a proxy for oncologic surgical success because nodal yield has been shown to be an independent prognostic factor for overall survival, disease-specific survival, and disease-free survival. In our analysis, the robotic approach yielded a similar number of total lateral and pathological nodes compared to the open approach. Concurrently, no statistically significant difference was found in 6-month recurrence rates, even after stratification by pathology and robotic technique. Although the similar nodal yields and short-term recurrence rates may provide support for the oncologic integrity of robotic neck dissection, studies including measures of true oncologic outcomes, such as locoregional, disease-free, and overall survival, remain necessary.

Our analysis also shows noninferior complication rates. No statistically significant differences in chyle leak, hematoma, Horner’s syndrome, marginal mandibular nerve weakness, seroma, wound infection, or amount of surgical drainage were noted. Given the rarity of complications associated with either technique, individual studies may have had limited capacity to detect a significant difference on their own, which can be resolved by pooling outcomes for a random effects meta-analysis. Major complications were rare; one study experienced an inadvertent injury to the internal jugular vein in the robotic cohort, but this was repaired without the need for open conversion. Ji et al. found that transient marginal nerve palsy was higher in the robotic group (11% vs. 1.3% at a statistically significant rate, $P = 0.043$), but all of these patients recovered within 2 to 3 months. Technique-specific complications were also reported by a number of studies. For example, recurrent laryngeal nerve palsy and hypocalcemia were reported by five transaxillary studies and found to have no significant differences between robotic and open groups. However, no meta-analysis was performed on these complications because they were related to the thyroidectomy portion of the procedures instead of the neck dissection. Other perioperative outcomes related to neck dissections, such as shoulder dysfunction, ear lobule numbness, skin flap necrosis, and pain, were described by a minority of included studies and found to be statistically similar between robotic and conventional approaches.

One of the primary benefits of robotic and laparoscopic surgeries is the capacity to perform minimally invasive interventions, which has been used to achieve superior cosmetic outcomes in the fields of general, plastic, and reconstructive surgery. A recent systematic review concluded that robotic-assisted laparoscopic radical prostatectomy (RALP) offers equivalent perioperative outcomes to open approaches with superior cosmesis. This conclusion is supported by the current study, which found no significant difference in subjective cosmesis scores between robotic and conventional approaches.

Random effects analysis

Fig. 4. Forest plot showing random effects meta-analysis for subjective cosmesis score. The pooled standardized mean difference (diamond) demonstrates a large and statistically significant effect in favor of robotic versus conventional approaches. When stratified by robotic techniques, the improved cosmesis provided by the retroauricular approach does not reach statistical significance. However, it remains statistically significant in the transaxillary approach. CI = confidence interval; RA/MFL = retroauricular/modified facelift; TA = transaxillary.
urologic, and gynecologic surgeries.\textsuperscript{1-6} Given the conspicuous location of traditional thyroidectomy and cervical lymphadenectomy incisions, cosmesis can be a priority for many head and neck cancer patients.\textsuperscript{54} The robotic techniques examined in our study placed the incision in a modified facelift, retroauricular, or transaxillary location, away from the center of the surgical field. This was associated with two primary advantages: 1) a less visible incision, and 2) an incision distant from the site of the postoperative radiation field. In accordance with the cosmetic outcomes reported in other surgical fields,\textsuperscript{1-6,55} scar satisfaction scores in the included studies were higher in robotic cohorts. This was confirmed with random effects meta-analysis, which showed a large effect size in favor of robotic neck dissection. When studies using retroauricular techniques were analyzed separately, the statistical significance of this difference disappeared, likely due to the limited number of pooled patients from these studies.

In the majority of included studies, scar satisfaction was evaluated at 3 months postoperatively. It is notable that some patients may have undergone adjuvant intervention at this time, which can interfere with appropriate wound healing. Ideally, follow-up cosmetic evaluation should take place at 1 year given that scar maturation occurs between 6 and 18 months, with complete healing and remodeling around 8 to 12 months.\textsuperscript{56,57} Despite the relatively short assessment period, two studies in the analysis evaluated scars at 6 months and 1 year. Lee et al.\textsuperscript{35} evaluated scar satisfaction at 6 months postoperatively, whereas Ji et al.\textsuperscript{27} followed cosmetic outcomes up to 1 year. Both studies continued to observe a significant difference in subjective satisfaction at these extended follow-up periods (\(P < 0.001\) and \(P = 0.006\), respectively)\textsuperscript{27,32}; however, these findings with regard to cosmesis must be interpreted with caution because the cosmetic scores in all the studies were administered as nonvalidated questionnaires to a cohort of patients who may have preemptively decided to undergo a more expensive, robotic-assisted procedure. Additionally, none of the studies clarified whether physician or patient provided the evaluation for cosmesis.\textsuperscript{28-32,35,36} The use of validated scar assessment tools at appropriate follow-up times would be important in making any definitive conclusions about cosmesis after robotic surgery.\textsuperscript{38-40}

The primary disadvantage associated with robotic neck dissection was increased operative duration. The extensive elevation of cervical skin flaps, docking, repositioning of the robot, and high learning curve combined to reduce operative time efficiency.\textsuperscript{6,61-63} All studies included in our meta-analysis demonstrated statistically significant differences in operative times in favor of the open approach, which was supported by a large effect size when the outcome was pooled. Many other articles have demonstrated that longer operative times increase the risk of perioperative complications\textsuperscript{59,64} and overall cost of care,\textsuperscript{65} making this an inherent obstacle to the adoption of robotic neck dissection in clinical practice. However, several studies\textsuperscript{29,30} in our meta-analysis did observe a trend toward a shorter duration of surgery with increased case load, indicating that this drawback may diminish with increased surgeon experience.

Only one of the included studies compared the actual financial costs of the procedures. Seup Kim et al.\textsuperscript{34} calculated direct and indirect costs related to performing a robotic total thyroidectomy with central compartment and lateral neck dissection and compared it to a conventional total thyroidectomy with MRND, demonstrating a statistically significant difference in costs favoring the conventional approach ($13,608 vs. $4,704, respectively, \(P < 0.001\)). However, because this study and the majority of others were performed in South Korea, it is important to keep in mind that extended hospital stays were often fully covered by the national health insurance system, which may have promoted extended postoperative hospitalization until stitches or closed-suction drains were removed.\textsuperscript{55} This was exemplified by the lack of a statistically significant difference in hospital days between robotic and open techniques in our meta-analysis. This particularity of the South Korean system may have artificially decreased the cost gap between robotic and conventional approaches because it would be reasonable to assume that longer operative times would confer longer postoperative stays. Studies in other countries, in the setting of different insurance systems, would be required to determine the true disparity in hospital stay and cost of care between these two groups.

There are several limitations to this review. The individual studies comprising the final data set were not randomized in design. In each of the studies, patients were given the option of choosing a conventional versus robotic surgical approach and were informed of the additional costs associated with the latter procedure. This translated into a high degree of selection bias, which may result in skewed distributions of complications rates, preoperative oncologic parameters, and postoperative adjuvant interventions. Furthermore, 10 of 11 studies were conducted in South Korea by six distinct group of researchers. This means that our results may not be broadly generalizable to other geographic locations, significantly restricting the external validity of this study.\textsuperscript{66} Additionally, operative time measurements were not standardized or fully described, with durations measured from incision to specimen removal or from incision to wound closure.\textsuperscript{33,38} As stated in the aforementioned discussion, scar-scoring systems were study-dependent and not validated, which restricted appropriate interpretation of this outcome.

With regard to statistical analysis, one study reported median and interquartile range for descriptive statistics,\textsuperscript{68} and the values were converted to mean and SD through a formula developed by Wan et al.\textsuperscript{46} The artificial conversion may reduce the accuracy and reliability of the data contributed by this study. Two studies\textsuperscript{10,32} were also noted to have potential overlap of patients who underwent conventional neck dissection, although their robotic study populations were distinct.

Yet another limitation is the diversity of robotic approaches studied. Our meta-analysis placed no restriction on the robotic approach to the lateral neck. Robotic approaches ranged from truly minimally invasive, such as in the bilateral axillo-breast approach, to what may be considered less invasive approaches, such as the
transaxillary and/or retroauricular approach. To mitigate this heterogeneity, we classified the robotic approach into two major techniques related to the placement of the incisions and performed our analyses together and separately based on the robotic technique. By coincidence, this also served to stratify the analysis based on pathology because all patients with PTC underwent a transaxillary approach, whereas patients with other pathologies underwent a retroauricular approach. In addition, our study combined all types of neck dissection for both elective and therapeutic indications, without any attempt at further stratification into SND versus MRND.

The variability among studies was reflected in the high heterogeneity parameter of certain outcomes, especially when analyzing cosmesis and operating room time. Although the I² was elevated among these outcomes, this was partially addressed by a random effects model, which assumes that a distribution of effect sizes exists. Moreover, the random effects model gives greater relative weight to small studies (in comparison to fixed effects analysis) and has wider 95% CIs. As shown in the results, despite the predilection for wider 95% CI with random effects, the majority of outcomes were associated with narrow 95% CI, which indicated adequate statistical power. The inherent biases present in each of the studies were somewhat tempered by their quality because most of them were prospective in nature and had favorable MINORS scores.

CONCLUSION

This is the first meta-analysis to examine whether there is evidence to support performing robotic cervical lymphadenectomy for head and neck cancer patients. Given the increasing utilization of robotic technology in head and neck surgery, a systematic evaluation of robotic approaches to the lateral neck is paramount in developing an evidence-based recommendation to patients in the clinical setting. Our study suggests that neck dissections performed robotically have similar nodal yields, recurrence rates, and complication rates but are also associated with significantly longer operative times compared to those performed conventionally. Additionally, whereas this meta-analysis shows that there may be a cosmetic benefit to utilizing a robotic approach, the failure to use validated cosmetic assessments in any of the studies make any conclusions regarding presence of cosmetic problematics. These findings demonstrate that there is limited evidence to justify utilizing the surgical robot to perform neck dissections for patients with head and neck malignancy. In addition, any future studies should be adequately powered to demonstrate oncologic noninferiority, cosmetic superiority, and limited difference in cost of care before any recommendation for the adoption of robotic cervical lymphadenectomy can be given.

Acknowledgments

Author Contributions: Daniel C. Sukato, lead author, study design and conception, data collection, data interpretation, and final approval; Daniel P. Ballard, contributing author, study design, data collection, data interpretation, and final approval; Jason M. Abramowitz, contributing author, study design, data collection, data interpretation, and final approval; Richard M. Rosenfeld, study design, statistical analysis, data interpretation, manuscript editor, and final approval; Stefan Mlot, senior author, study design and conception, data interpretation, manuscript editor, and final approval.

BIBLIOGRAPHY

6. Tsung A, Geller DA, Sukato DC, et al. Robotic versus laparoscopic hepatec-
27. Ji YB, Song CM, Bang HS, Park HJ, Lee JY, Taek K. Functional and cos-
Laryngoscope 129: July 2019

Sukato et al.: Robotic Versus Open Neck Dissection


