Effect of intraoperative neuromonitoring on recurrent laryngeal nerve palsy rates after thyroid surgery—A meta-analysis

Shixing Zheng, Zhiwen Xu*, Yuanyuan Wei, Manli Zeng, Jinnian He

Department of Otolaryngology Head and Neck Surgery, First Affiliated Hospital of Guangxi Medical University, Nanning, Guangxi 530021, China

Received 24 August 2011; received in revised form 1 March 2012; accepted 6 March 2012

Background/Purpose: Though intraoperative nerve monitoring (IONM) during thyroid surgery has gained universal acceptance for localizing and identifying the recurrent laryngeal nerve (RLN), its role in reducing the rate of RLN injury remains controversial. In order to assess the effect of IONM during thyroid surgery, its value in reducing the incidence of RLN palsy was systematically evaluated.

Methods: Studies were evaluated for inclusion in this analysis by researching PubMed, Embase, the Cochrane Central Register of Controlled Trials, and the references of included studies. The initial screening of article titles and abstracts was independently performed by five reviewers based on the research protocol criteria. Each article was then read in detail and discussed before inclusion in the meta-analysis. Data were independently extracted, including the level of evidence, number of at-risk nerves, allocation method, baseline equivalence between groups, definitions of transient and permanent vocal fold palsy, systematic application of electrodes, etc. The meta-analysis was then performed. Odds ratios were pooled using a random effects model.

Results: Five randomized clinical trials and 12 comparative trials evaluating 36,487 at-risk nerves were included. Statistically significant differences in terms of total recurrent laryngeal nerve palsy (3.37% with intraoperative nerve monitoring [IONM] vs. 3.76% without IONM [OR: 0.74; 95% confidence interval [CI]: 0.59–0.92]) and transient recurrent laryngeal nerve palsy (2.56% with IONM vs. 2.71% without IONM [OR: 0.80; 95% CI: 0.65–0.99]) were identified. The persistent incidence of recurrent laryngeal nerve palsy was 0.78% for IONM versus 0.96% for nerve identification alone (OR: 0.80; 95% CI: 0.62–1.03).

Conclusion: Based on this meta-analysis, statistically significant differences were determined in terms of the incidences of total and transient recurrent laryngeal nerve palsy after using IONM versus...
Introduction

Recurrent laryngeal nerve (RLN) injury has been a frequent source of malpractice litigation following thyroid surgery over the last 20 years. The average incidences of permanent and temporary RLN palsy after thyroid operations are high (2.3% and 9.8%, respectively), as verified by systematic and rigorous postoperative laryngoscopy. Most surgeons apply their best efforts to prevent this considerable postoperative complication, which involves hoarseness, impaired vocal register, dysphonia, dysphagia, and aspiration dyspnea. The rate of RLN palsy depends on the type of disease (benign or malignant), the extent of thyroid resection (subtotal or total thyroidectomy), the type of resection (first time or reoperation), the surgical device used (with or without routine RLN identification), and the training and/or experience of the surgeon. Surgical exposure and anatomic identification of the RLN during thyroid operations has been shown to provide the best way of keeping the incidence of nerve injury to a minimum. However, even experienced surgeons can inadvertently injure the nerve due to variability in RLN anatomy and difficulties in nerve identification that can occur under challenging conditions.

Intraoperative nerve monitoring (IONM) during thyroid surgery has gained universal acceptance as an adjunct, not only for localizing and identifying the RLN but also as a way to predict vocal cord function and clarify the mechanisms of RLN injury. The rates of the use of monitoring alone on the rate of RLN palsy in thyroid surgery were also excluded from this review. In addition, a reference- and related-article search was performed. The articles that were generated in each step of the search were then aggregated to produce an article list for consideration.

Five investigators independently screened the article titles and abstracts in order to evaluate relevant articles, which included all randomized controlled trials (RCTs) and cohort studies related to RLN monitoring during thyroid surgery. The investigators then obtained the full texts that were deemed relevant to the targeted research purposes for further analysis (Fig. 1). Additional studies were identified by scanning the references of the initial studies. Articles were then excluded from the final selected cohort if they did not meet the original parameters of the search criteria.

Data extraction and quality assessment

All articles that met the eligibility criteria were independently reviewed by the three investigators for data extraction and quality assessment. Data extracted from each trial included the level of evidence (LOE), number of nerves at risk (NAR), allocation method, baseline equivalence between groups, the definitions of transient and permanent vocal fold palsy, systematic application of electrodes, country, and dates of during which the studies were conducted. Primary outcome variables included total, transient, and persistent RLN palsy. Total RLN palsy was defined as transient RLN palsy in combination with permanent RLN palsy.

Statistical analysis

In order to decrease heterogeneity across study interventions, statistical analyses were performed after all of the authors had critically evaluated the articles and decided on the final list of included articles. An odds ratio (OR) of 0.5 was chosen to represent an estimate that could be considered clinically significant and might change the clinical decision. The results were pooled using RevMan 5.0.23 meta-analysis software using the random effects
Mantel-Haenszel model. The OR was used to reflect the association across studies. Statistical heterogeneity was calculated using Cochrane Q and the Tau² statistics. The potential sources of heterogeneity considered in the study protocol included trial quality, the definitions of transient and permanent vocal fold palsy, and the risk level of RLN. Sensitivity analyses were also performed to assess whether the principal conclusions were affected by basic assumptions. A funnel plot was constructed to evaluate data symmetry and publication bias. Forest plots were generated to represent the OR with the corresponding 95% confidence intervals (CI) across the included studies.

Results

Study selection and study characteristics

The electronic searches and review of the references yielded a total of 206 potentially relevant articles. Fig. 1 shows the flowchart of the inclusion and exclusion of the potential studies. During the review of the title, abstract, and full text, an additional 192 articles were excluded, resulting in 14 articles that were finally included in this meta-analysis. Two studies were multi-institutional. These studies collectively enrolled a total of 36,487 NARs that had undergone IONM. According to the Oxford Centre for Evidence-Based Medicine (OCEBM),²⁶ two and 12 articles were graded as 1b and 2b, respectively (Table 1).²⁶ Additional details are summarized in Table 2. Five studies were conducted in the United States, four in Germany, and five in other countries (China, Italy, Poland, Brazil, and Turkey). These studies were published between 1992 and 2009.

The patient-allocation method for comparative studies included random, consecutive allocation, surgeon preference, and equipment availability. The timing of the initial postoperative laryngoscopy varied from immediately postoperative to within 2 weeks of surgery. Additionally, persistent vocal fold palsy was variously defined as 3 months, 6 months, or even 12 months. Most studies did not specify the primary surgeon but mentioned the exact ratios of low- and high-volume surgeons.

Risk of bias in the included studies

The risk of bias in the included studies is summarized in the risk-of-bias graph (Fig. 2) and risk-of-bias summary (Fig. 3). Only two of the 14 trials demonstrated adequate sequence generation.⁵,¹¹ None of the other 10 trials reported the method used to generate the allocation sequence. Only one trial performed allocation concealment.⁵ Only one trials reported patient blinding.⁵ None of the other trials reported patient or assessor blinding, which are theoretically feasible in this comparison. Four trials reported post-randomization drop-outs. However, one trial excluded some patients from the analysis without interpretation.¹⁵ Therefore, this trial was considered to be at high risk of bias from incomplete outcome data. The other studies did not report the participant flow and we were incapable of assessing bias from incomplete outcome data. Four of the trials did not report all of the important outcomes (e.g., primary outcomes), and we were unable to acquire the protocols of these trials. Therefore, we consider these four trials to be at high risk of bias from selective outcome reporting. Two trials did not report the participant characteristics.²¹,²² Nine trials had adequately matched participants to two groups and were free from baseline imbalance bias. However, three trials did not adequately match participants and were at risk of baseline imbalance bias.¹²,¹³,¹⁶ The source of funding was described in seven trials. Terris DJ served as a consultant for thyroid

---

Table 1

<table>
<thead>
<tr>
<th>GOR</th>
<th>LOE</th>
<th>Therapy/prevention/etiology/harm</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1a</td>
<td>SR (with homogeneity) of RCTs</td>
</tr>
<tr>
<td></td>
<td>1b</td>
<td>Individual RCT (with narrow confidence interval)</td>
</tr>
<tr>
<td></td>
<td>1c</td>
<td>All or none</td>
</tr>
<tr>
<td>B</td>
<td>2a</td>
<td>SR (with homogeneity) of cohort studies</td>
</tr>
<tr>
<td></td>
<td>2b</td>
<td>Individual cohort study (including low quality RCT; e.g., &lt; 80% follow-up)</td>
</tr>
<tr>
<td></td>
<td>2c</td>
<td>“Outcomes” research; ecological studies</td>
</tr>
<tr>
<td></td>
<td>3a</td>
<td>SR (with homogeneity) of case-control studies</td>
</tr>
<tr>
<td></td>
<td>3b</td>
<td>Individual case-control study</td>
</tr>
<tr>
<td>C</td>
<td>4</td>
<td>Case series (and poor quality cohort and case-control studies)</td>
</tr>
<tr>
<td>D</td>
<td>5</td>
<td>Expert opinion without explicit critical appraisal, based on physiology, bench research or “first principles”</td>
</tr>
</tbody>
</table>

GOR = grades of recommendation; LOE = level of evidence; RCT = randomized controlled trial; SR = systematic review.
instrument development for Medtronic Xomed and directed thyroid courses sponsored by Ethicon Endosurgery. Therefore, this trial was at risk of source-of-funding bias. The sources of funding of the other six trials were medical foundations, and we consider these trials to be free from risk of source-of-funding bias. The source-of-funding bias was unclear in the remaining seven trials, which did not declare the sources of funding.

Internal analyses and pooled analyses

The Cochrane Q and the Tau² statistics (using a significance threshold of \( p < 0.05 \)) did not reveal significant statistical heterogeneity among studies. The funnel plots indicated a relatively perfect symmetrical distribution regarding the rates of transient and persistent RLN palsy. However, mild asymmetry was observed in the funnel plot of the total RLN palsy rate (Fig. 4).

Fig. 5 shows the forest plots of all of the included surgeries. The rate of total RLN palsy of the IONM group and the without IONM group are 3.37% and 3.76%, respectively. The rate of transient RLN nerve palsy (2.56% with IONM vs. 2.71% without IONM) was identified. Additionally, the persistent incidence of recurrent laryngeal nerve palsy was 0.78% for IONM versus 0.96% for nerve identification alone. The pooled adjusted ORs for the rates of total, transient, and persistent RLN palsy, respectively, were 0.74 (0.59–0.92), 0.80 (0.65–0.99), and 0.80 (0.62–1.03). When the total and transient RLN palsy rates were compared, there was a significant difference between the two groups (\( p = 0.007 \) and 0.04, respectively). However, comparison of the persistent RLN palsy rates between groups revealed no statistically significant differences (\( p = 0.08 \)).

Sensitivity analyses

It is acknowledged that nerve damage is much higher in reoperative cases of thyroid and parathyroid surgery. A study conducted by Yarbrough et al predominantly focused on reoperative thyroid and parathyroid surgery. It is interesting to speculate whether the results would be the same if this study was excluded. To test the sensitivity of this meta-analysis, all analyses were repeated. The combined ORs did not change substantially; ORs of 0.72 (0.57–0.91) for the total RLN palsy rate, 0.78 (0.62–0.97) for the transient RLN palsy rate, and 0.80 (0.62–1.03) for the permanent RLN palsy rate were returned upon reanalysis.

Discussion

Although different types of IONM have been adopted, many potential benefits and pitfalls have also been reported. Nerve identification, which helps to differentiate RLN from nonneural tissue, is a primary benefit of RLN monitoring. Another benefit of IONM is dynamic and
continuous audible and visual feedback of mechanically evoked potentials so that the surgical methods can be modified when something potentially harmful is being done to the RLN, especially during a planned exploration procedure in a patient at increased risk of bilateral nerve injury.5,28 Moreover, by printing the electromyographic signal of the evoked potentials, RLN monitoring allows nerve function documentation both before and after thyroid resection, which is of great importance in a society with an increasing number of litigations.31,32 However, there are several criticisms of the routine use of IONM, which are focused mainly on the increased cost of the equipment (i.e., the cost of the special endotracheal tube and the additional technical staff who intraoperatively monitor the RLN).11,15,21 Another criticism is that it is time consuming.23,33 The disadvantages associated with IONM include false-negative and false-positive signals,28,29,34 which should not be neglected, as false-negative signals can actually increase the risk of RLN injury if surgeons exclusively rely on the monitoring system. Despite the increasing popularity of the adoption of IONM, its role in reducing the incidence of postoperative RLN palsy remains doubtful. Additional evidence-based reviews of the clinical utility of IONM are needed.

In this meta-analysis, when comparing the "with IONM" and "without IONM" groups, a significant difference was observed in the reduction of the rates of total and transient RLN palsies. But the reduction of the rates of persistent RLN palsies was not observed in adequately powered sample sizes. However, the sensitivity analysis revealed that the pooled analysis of the rate of transient RLN palsy was hypersensitive and unstable. Thus, we may draw conclusions that IONM plays an important role in reducing the rate of total RLN palsy but not the rate of persistent RLN palsy. It is prudent to conclude that IONM effectively decreases the rate of transient RLN palsy.

Barczyński et al5 performed a medium-sized, single-center, randomized, prospective study that compared the frequency of transient RLN paresis after surgery between RLN visualization alone and with IONM. This study demonstrated that nerve monitoring decreases the incidence of transient RLN paresis but not permanent RLN paresis in comparison with visualization alone, particularly in high-risk patients. To date, although this is the highest level study performed, the results and conclusions are still controversial. Nonrandomized comparative trials have returned fairly accurate results, and their utilization has been advocated for special conditions in which randomized trials are difficult to perform.13 Until further large-scale randomized controlled trials (RCTs) can be performed, data from published nonrandomized studies can be used for such comparisons, as shown above.25,39 Due to insufficient numbers, most of the included studies failed to demonstrate that a statistically significant difference
Figure 4 Funnel plots depicting the odds ratios (ORs) with 95% confidence intervals (CI) and standard errors (SE) for total (A), transient (B), and persistent (C) recurrent laryngeal nerve palsies in the included studies. Total recurrent laryngeal nerve palsy is defined as transient recurrent laryngeal nerve palsy in combination with permanent recurrent laryngeal nerve palsy.

Figure 5 Forest plots showing the distributions of (A) total recurrent laryngeal nerve palsy, (B) transient recurrent laryngeal nerve palsy, and (C) persistent recurrent laryngeal nerve palsy. Total recurrent laryngeal nerve palsy is defined as transient recurrent laryngeal nerve palsy in combination with permanent recurrent laryngeal nerve palsy. CI = confidence interval.
between the two groups of unintentional RLN injury actually existed. Large prospective randomized trials that address these issues are rare because of the large numbers of patients required (> 7000 per arm) to indicate a significant difference 0.2% (from 0.3–0.5% per NAR). Eisele noted that in order to show a reduction in the RLN paralysis rate from 2% to 1% per NAR, a study group of approximately 1000 patients would be needed. Further randomized controlled trials are needed to confirm these results. Unfortunately, a prospective, multicenter, randomized, control study is difficult to perform because many patients and surgeons expect nerve monitoring to supply additional benefits, thus limiting randomization to many patients and surgeons expect nerve monitoring to supply additional benefits, thus limiting randomization to a non-IONM control group. Most articles neglected to describe or even mention the circumstances of any unintentional RLN injuries, which made this comparison difficult to assess. A study based on a systematic appraisal of the literature using evidence-based criteria reached the conclusion that apart from navigating the surgeon through challenging anatomies, IONM may be useful as a routine adjunct to the gold standard of visual nerve identification. However, a meta-analysis performed by Thomas, which included one randomized clinical trial, seven comparative trials, and 34 case series, demonstrated no statistically significant differences in the rates of RLN palsy after using IONM versus recurrent laryngeal nerve identification alone during thyroidectomy. Obviously, this conclusion is contradictory to our results. In our meta-analysis, two randomized clinical trials and 12 comparative trials were included without a case series. As Thomas discussed in his study, case series involve biases, including inherent selection bias, publication bias, allocation bias, measurement bias, and intervention bias. These biases may contribute to these diverse outcomes. In the meta-analysis performed by Thomas, subgroup analyses were also performed, which showed that none of these comparisons demonstrated statistically significant differences. In fact, the individual patient dates (IPDs) of the studies included in the Thomas study and ours were imperfect. It is possible that differences between subgroups might disappear or become significant with detailed IPDs.

The large number of studies suggesting the utility of IONM and its current clinical use indicate the urgent need for a well-designed study that assesses its accuracy and clinical value. The current lack of sufficient evidence is due to the absence of high-quality evidence-based studies that have investigated these applications. The studies that are currently available on this issue include biases and lack standardized methodology. Randomization is a critical feature of study design that should be included in future studies. To determine the utility of IONM in the prevention of RLN palsy, clinical investigators need to randomly bring patients into experimental and control groups. Additional evidence-based studies are recommended to verify the value of IONM, the optimal electrode type for specific situations, and the relationship with prevention outcomes. Post-operative laryngoscopic examination should be performed within 24 hours after surgery and then periodically until paralysis resolves. We recommended a definition of persistent vocal fold palsy as palsy lasting no less than 12 months. Complete demographics (e.g., age, gender, pathology, surgery) and subgroup stratification analyses should also be included.

This study has several limitations that merit further discussion. Although we have made attempts to be as inclusive and as transparent as possible when generating our search strategy, there may be articles that were inadvertently filtered out by our search parameters. In addition, excluding non-English articles does introduce potential bias to our review. Additionally, we limited our review to published articles and did not perform a deliberate search for unpublished data. It is well known that heterogeneity is an inherent limitation to meta-analysis. This was addressed in our study using standard meta-analytic techniques, including the Cochrane Q and Tau² statistics, random effects analysis, sensitivity analysis, and standardized article selection criteria. Subgroup analysis was not performed because of the insufficient data supplied in the included studies.

Based on this meta-analysis, we were able to identify statistically significant differences in the incidences of total and transient RLN palsy after using IONM versus RLN identification alone during routine thyroidectomy, though a hypersensitivity was revealed in the pooled analysis of the transient RLN palsy rate. However, no statistically significant difference was identified in the incidence of persistent RLN palsy between the IONM group and without IONM group. Hopefully, this review will lead to more high-level studies, and, in the future, more can be stated regarding the incidence of recurrent laryngeal nerve palsy with and without IONM during thyroidectomy.

Disclosures
The authors have no competing interests or sponsorships to declare.
## Appendix 1

<table>
<thead>
<tr>
<th>Data bases</th>
<th>Search strategy</th>
<th>Results</th>
</tr>
</thead>
</table>
| **PUBMED (1950 — April, 2011)** | 1. thyroidectomy [mh]  
2. (Thyroid Gland [mh] OR thyroid [tw]) AND Surgical Procedures, Operative [mh]  
3. OR/1-2  
5. vocal cord paralysis [mh]  
6. (Vocal fold* [tw] OR Vocal cord* [tw]) AND (immobility [tw] OR palsy [tw])  
7. (Inferior Laryngeal Nerve* [tw] OR Recurrent laryngeal nerve* [tw]) AND palsy [tw]  
8. OR/5-7  
9. monitoring, intraoperative [mh]  
11. Electrophysiologic monitoring [tw]  
12. laryngeal electromyography [tw]  
13. or/9-12  
14. 3 and 8 and 13 | 119 |
| **Cochrane Central Register of Controlled Trials (CENTRAL, The Cochrane Library, Issue 2 of 4, Apr 2011)** | #1. MeSH descriptor Thyroid Gland explode all trees  
#2. thyroid*  
#3. MeSH descriptor Surgical Procedures, Operative explode all trees  
#4. (#1 OR #2) AND #3  
#5. MeSH descriptor Thyroidectomy explode all trees  
#6. thyroidectomy*  
#7. (#5 OR #6)  
#8. MeSH descriptor Vocal Cord Paralysis explode all trees  
#9. (Vocal fold*) or (Vocal cord*)  
#10. (immobilit*) or (pals*)  
#11. #8 OR (#9 AND #10)  
#12. MeSH descriptor Recurrent Laryngeal Nerve explode all trees  
#13. (Inferior Laryngeal Nerve*)  
#14. (Recurrent laryngeal nerve*)  
#15. #12 OR #13 OR #14  
#16. MeSH descriptor Monitoring, Intraoperative explode all trees  
#17. (neuromonitoring) or neurostimulation or nerve Monitoring or (monitoring)  
#18. (Intraoperative) or (Intra-operative)  
#19. #16 OR (#17 AND #18)  
#20. Electrophysiologic monitoring  
#21. (laryngeal electromyography)  
#22. #19 OR #20 OR #21  
#23. (#4 OR #7) AND #11 AND #15 AND #22 | 4 |
| **EMBASE (1974 — April, 2011)** | #1. 'thyroid gland'/exp  
#2. thyroid*  
#3. 'surgery'/exp  
#4. #1 OR #2  
#5. #3 AND #4  
#6. thyroidectomy*  
#7. 'thyroidectomy'/exp  
#8. #6 OR #7  
#9. #5 OR #8  
#10. vocal AND fold*  
#11. vocal AND cord*  
#12. #10 OR #11  
#13. immobilit* | 83 |
References


16. Brauckhoff M, Gimm O, Thanh PN. First experiences in intraoperative neurostimulation of the recurrent laryngeal nerve


