Cochrane Library

Cochrane Database of Systematic Reviews

Intraoperative neuromonitoring versus visual nerve identification for prevention of recurrent laryngeal nerve injury in adults undergoing thyroid surgery (Review)

Cirocchi R, Arezzo A, D'Andrea V, Abraha I, Popivanov GI, Avenia N, Gerardi C, Henry BM, Randolph J, Barczyński M

Cirocchi R, Arezzo A, D'Andrea V, Abraha I, Popivanov GI, Avenia N, Gerardi C, Henry BM, Randolph J, Barczyński M.
Intraoperative neuromonitoring versus visual nerve identification for prevention of recurrent laryngeal nerve injury in adults undergoing thyroid surgery.
Cochrane Database of Systematic Reviews 2019, Issue 1. Art. No.: CD012483.
DOI: 10.1002/14651858.CD012483.pub2.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

HEADER . . . 1
ABSTRACT . . . 1
PLAIN LANGUAGE SUMMARY . . . 2
SUMMARY OF FINDINGS FOR THE MAIN COMPARISON . . . 4
BACKGROUND . . . 7
Figure 1. . . . 8
Figure 2. . . . 9
Figure 3. . . . 10
OBJECTIVES . . . 11
METHODS . . . 11
RESULTS . . . 16
Figure 4. . . . 17
Figure 5. . . . 19
Figure 6. . . . 20
DISCUSSION . . . 22
AUTHORS' CONCLUSIONS . . . 23
ACKNOWLEDGEMENTS . . . 24
REFERENCES . . . 24
CHARACTERISTICS OF STUDIES . . . 29
DATA AND ANALYSES . . . 47
Analysis 1.1. Comparison 1 Intraoperative neuromonitoring versus visual nerve identification only, Outcome 1 Permanent RILN palsy. . . . 47
Analysis 1.2. Comparison 1 Intraoperative neuromonitoring versus visual nerve identification only, Outcome 2 Transient RILN palsy. . . . 48
Analysis 1.3. Comparison 1 Intraoperative neuromonitoring versus visual nerve identification only, Outcome 3 Adverse events other than RILN palsy. . . . 49
Analysis 1.4. Comparison 1 Intraoperative neuromonitoring versus visual nerve identification only, Outcome 4 Operative time. . . . 50
Analysis 1.5. Comparison 1 Intraoperative neuromonitoring versus visual nerve identification only, Outcome 5 All-cause mortality. . . . 51
ADDITIONAL TABLES . . . 51
APPENDICES . . . 55
CONTRIBUTIONS OF AUTHORS . . . 79
DECLARATIONS OF INTEREST . . . 80
SOURCES OF SUPPORT . . . 80
DIFFERENCES BETWEEN PROTOCOL AND REVIEW . . . 81
NOTES . . . 81

[Intervention Review]

Intraoperative neuromonitoring versus visual nerve identification for prevention of recurrent laryngeal nerve injury in adults undergoing thyroid surgery

Roberto Cirocchi[1], Alberto Arezzo[2], Vito D'Andrea[3], Iosief Abraha[4], Georgi I Popivanov[5], Nicola Avenia[6], Chiara Gerardi[7], Brandon Michael Henry[8], Justus Randolph[9], Marcin Barczyński[10]

[1]Department of General Surgery, University of Perugia, Terni, Italy. [2]Department of Surgical Sciences, University of Torino, Turin, Italy. [3]Department of Surgical Sciences, Sapienza University of Rome, Rome, Italy. [4]Health Planning Service, Regional Health Authority of Umbria, Perugia, Italy. [5]Department of Surgery, Medical Military Academy of Sofia, Sofia, Bulgaria. [6]Department of Surgical Sciences, University of Perugia, Perugia, Italy. [7]IRCCS Istituto di Ricerche Farmacologiche Mario Negri, Milan, Italy. [8]Division of Cardiology, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio, USA. [9]Georgia Baptist College of Nursing, Mercer University, Atlanta, GA, USA. [10]Department of Endocrine Surgery, Third Chair of General Surgery, Jagiellonian University, Medical College, Krakow, Poland

Contact address: Roberto Cirocchi, Department of General Surgery, University of Perugia, Terni, 05100, Italy. cirocchiroberto@yahoo.it.

Editorial group: Cochrane Metabolic and Endocrine Disorders Group.
Publication status and date: New, published in Issue 1, 2019.

Citation: Cirocchi R, Arezzo A, D'Andrea V, Abraha I, Popivanov GI, Avenia N, Gerardi C, Henry BM, Randolph J, Barczyński M. Intraoperative neuromonitoring versus visual nerve identification for prevention of recurrent laryngeal nerve injury in adults undergoing thyroid surgery. *Cochrane Database of Systematic Reviews* 2019, Issue 1. Art. No.: CD012483. DOI: 10.1002/14651858.CD012483.pub2.

ABSTRACT

Background

Injuries to the recurrent inferior laryngeal nerve (RILN) remain one of the major post-operative complications after thyroid and parathyroid surgery. Damage to this nerve can result in a temporary or permanent palsy, which is associated with vocal cord paresis or paralysis. Visual identification of the RILN is a common procedure to prevent nerve injury during thyroid and parathyroid surgery. Recently, intraoperative neuromonitoring (IONM) has been introduced in order to facilitate the localisation of the nerves and to prevent their injury during surgery. IONM permits nerve identification using an electrode, where, in order to measure the nerve response, the electric field is converted to an acoustic signal.

Objectives

To assess the effects of IONM versus visual nerve identification for the prevention of RILN injury in adults undergoing thyroid surgery.

Search methods

We searched CENTRAL, MEDLINE, Embase, ICTRP Search Portal and ClinicalTrials.gov. The date of the last search of all databases was 21 August 2018. We did not apply any language restrictions.

Selection criteria

We included randomised controlled trials (RCTs) comparing IONM nerve identification plus visual nerve identification versus visual nerve identification alone for prevention of RILN injury in adults undergoing thyroid surgery

Data collection and analysis

Two review authors independently screened titles and abstracts for relevance. One review author carried out screening for inclusion, data extraction and 'Risk of bias' assessment and a second review author checked them. For dichotomous outcomes, we calculated risk ratios (RRs) with 95% confidence intervals (CIs). For continuous outcomes, we calculated mean differences (MDs) with 95% CIs. We assessed trials for certainty of the evidence using the GRADE instrument.

Main results

Five RCTs with 1558 participants (781 participants were randomly assigned to IONM and 777 to visual nerve identification only) met the inclusion criteria; two trials were performed in Poland and one trial each was performed in China, Korea and Turkey. Inclusion and exclusion criteria differed among trials: previous thyroid or parathyroid surgery was an exclusion criterion in three trials. In contrast, this was a specific inclusion criterion in another trial. Three trials had central neck compartment dissection or lateral neck dissection and Graves' disease as exclusion criteria. The mean duration of follow-up ranged from 6 to 12 months. The mean age of participants ranged between 41.7 years and 51.9 years.

There was no firm evidence of an advantage or disadvantage comparing IONM with visual nerve identification only for permanent RILN palsy (RR 0.77, 95% CI 0.33 to 1.77; P = 0.54; 4 trials; 2895 nerves at risk; very low-certainty evidence) or transient RILN palsy (RR 0.62, 95% CI 0.35 to 1.08; P = 0.09; 4 trials; 2895 nerves at risk; very low-certainty evidence). None of the trials reported health-related quality of life. Transient hypoparathyroidism as an adverse event was not substantially different between intervention and comparator groups (RR 1.25; 95% CI 0.45 to 3.47; P = 0.66; 2 trials; 286 participants; very low-certainty evidence). Operative time was comparable between IONM and visual nerve monitoring alone (MD 5.5 minutes, 95% CI −0.7 to 11.8; P = 0.08; 3 trials; 1251 participants; very low-certainty evidence). Three of five included trials provided data on all-cause mortality: no deaths were reported. None of the trials reported socioeconomic effects. The evidence reported in this review was mostly of very low certainty, particularly because of risk of bias, a high degree of imprecision due to wide confidence intervals and substantial between-study heterogeneity.

Authors' conclusions

Results from this systematic review and meta-analysis indicate that there is currently no conclusive evidence for the superiority or inferiority of IONM over visual nerve identification only on any of the outcomes measured. Well-designed, executed, analysed and reported RCTs with a larger number of participants and longer follow-up, employing the latest IONM technology and applying new surgical techniques are needed.

PLAIN LANGUAGE SUMMARY

Identification of nerves using an electrode compared with visual nerve identification for adults undergoing thyroid surgery

Review question

To assess the effects of intraoperative neuromonitoring compared with visual nerve identification for the prevention of recurrent laryngeal nerve injury during thyroid surgery in adults.

Background

Thyroidectomy is an operation that removes a part or all of the thyroid gland to cure benign disorders (for example multinodular goitre), or cancer. The recurrent laryngeal nerves are responsible for the movement of the vocal cords and they can easily be injured during thyroid surgery, resulting in one-sided or two-sided vocal cord paralysis leading to difficulty in speaking (dysphonia), breathing problems, or both. This, in turn may result in decreased health-related quality of life and may lead to permanent disability. Visual identification of recurrent laryngeal nerves during surgery has long been a standard procedure to prevent their injury. Recently, intraoperative neuromonitoring, which identifies nerves using an electrode, has been introduced to help the surgeon find and protect the recurrent laryngeal nerves.

Trial characteristics

We searched for randomised controlled trials (trials where the participants are randomly allocated to one, two, or more treatment arms), comparing intraoperative neuromonitoring plus visual nerve identification with visual nerve identification alone. We included only trials that reported data on participants older than 18 years who underwent thyroid surgery. We excluded trials with participants with previous neck surgery or repeated laryngeal nerve paralysis. We included five trials with a total of 1558 participants; 781 participants

were allocated to intraoperative neuromonitoring and 777 participants were allocated to visual nerve identification only. Two trials took place in Poland and one study each in Turkey, China and Korea. The average age of participants ranged between 41.7 years and 51.9 years.

This evidence is up to date as of 21 August 2018.

Key results

There was no firm evidence of an advantage or disadvantage comparing intraoperative neuromonitoring with visual nerve identification only for permanent or temporary recurrent laryngeal nerve paralysis, side effects and the time the operation lasted. Three of five included trials had data on all-cause mortality and reported no deaths. None of the trials reported health-related quality of life or socioeconomic effects (for example costs related to hospital stay). We need well-designed, executed, analysed and reported trials, with a larger number of participants and longer observation periods after surgery, using the latest intraoperative neuromonitoring technology and applying new surgical techniques.

Certainty of the evidence

We are very uncertain about the effects of intraoperative neuromonitoring compared with visual nerve identification for the prevention of recurrent laryngeal nerve injury during thyroid surgery. There were only a small number of trials, some systematic errors in the included trials and results were imprecise.

SUMMARY OF FINDINGS FOR THE MAIN COMPARISON [Explanation]

Intraoperative neuromonitoring compared to visual nerve identification only

Patients: adults undergoing thyroid surgery
Setting: inpatients
Intervention: intraoperative neuromonitoring
Comparison: visual nerve identification only

Outcomes	Risk with visual nerve identification only	Risk with intraoperative neuromonitoring	Relative effect (95% CI)	№ of participants/ nerves at risk (trials)	Certainty of the evidence (GRADE)	Comments
Permanent RILN palsy (nerves) Definition: injury detected clinically, by laryngoscopy or both, in which the motility of the vocal cords did not recover within 6 months after surgery Follow-up: 6-12 months	**9 per 1000**	**7 per 1000** (3 to 16)	**RR 0.77** (0.33 to 1.77)	2895 (4)	⊕◯◯◯ **Very low**[a]	Numbers refer to 'nerves at risk'; CI probably wider because of clustered data The 95% prediction interval ranged between 0.12 and 4.79
Transient RILN palsy (nerves) Definition: injury detected clinically, by laryngoscopy or both, in which the motility of the vocal cords recovered within 6 months after surgery Follow-up: 6-12 months	**36 per 1000**	**22 per 1000** (13 to 39)	**RR 0.62** (0.35 to 1.08)	2895 (4)	⊕◯◯◯ **Very low**[b]	Numbers refer to 'nerves at risk'; CI probably wider because of clustered data The 95% prediction interval ranged between 0.12 and 3.11
Health-related quality of life	Not reported					

Adverse events other than RILN palsy (participants) Definition: transient hypoparathyroidism Follow-up: 6-12 months	**122 per 1000**	**153 per 1000** (55 to 424)	**RR 1.25** (0.45 to 3.47)	286 (2)	⊕◯◯◯ **Very low**[c]	-
Operative time (min) Definition: time from the first skin incision to skin closure Follow-up: 6-12 months	The mean operative time ranged across control groups from 82.4 min to 274.2 min	The mean operative time in the intervention groups was **5.5 min longer** (0.7 min shorter to 11.8 min longer)	-	1251 (3)	⊕◯◯◯ **Very low**[d]	The 95% prediction interval ranged between −60.6 min and 71.7 min
All-cause mortality (nerves) Follow-up: 6-12 months	See comment			1438 (3)	⊕⊕⊕◯ **Moderate**[e]	3 of 5 trials provided data on all-cause mortality, no deaths were reported
Socioeconomic effects	Not reported					

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: confidence interval; **RILN:** recurrent inferior laryngeal nerve; **RR:** risk ratio

GRADE Working Group grades of evidence
High certainty: We are very confident that the true effect lies close to that of the estimate of the effect
Moderate certainty: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low certainty: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
Very low certainty: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

[a]Downgraded by one level because of performance bias, by one level because of inconsistency (no consistent direction of effects) and by one level because of imprecision (CI consistent with both benefit and harm, small number of trials, not a common event, i.e. ≤ 1/100) - see Appendix 15

[b]Downgraded by one level because of performance bias, by one level because of inconsistency (no consistent direction of effects) and by one level because of imprecision (CI consistent with both benefit and harm, small number of trials) - see Appendix 15

[c]Downgraded by one level because of inconsistency (no consistent direction of effects) and by two levels because of imprecision (CI consistent with both benefit and harm, small number of trials) - see Appendix 15

[d]Downgraded by one level because of performance bias, by one level because of inconsistency (no consistent direction of effects, wide 95% prediction interval) and by one level because of imprecision (CI consistent with both benefit and harm, small number of trials) - see Appendix 15

[e]Downgraded by one level because of imprecision (small number of trials) - see Appendix 15

BACKGROUND

Description of the condition

The visual identification of the recurrent inferior laryngeal nerve (RILN) is considered the safest method to prevent nerve injury during thyroid and parathyroid surgery (Deniwar 2015a). Generally, the rate of nerve injury is higher in cases of thyroid carcinoma, Flajani-Graves-Basedow disease, goitre, thyroid reoperation surgery, failure of nerve identification, and surgeons' inexperience (Calò 2014a). Intraoperative neuromonitoring (IONM) has been introduced in order to facilitate the localisation of the RILN and prevent its injury during surgery (Duclos 2011). A trial that included 686 participants demonstrated that use of IONM decreased the incidence of RILN palsy (from 7.6% to 4.7%; Duclos 2011). IONM was reported to reduce the prevalence of transient RILN injury (Barczyn ski 2009), and to increase surgeons' accuracy during nerve preparation, particularly during video-assisted thyroid surgery (Dionigi 2009).

When used by experienced thyroid surgeons (RILN injury rate of less than 1%), the IONM did not show a significant improvement in the postsurgery outcomes (Barczyn ski 2009). However, in procedures performed by low-volume surgeons (defined as surgeons that perform fewer than 25 thyroidectomies per year (Adam 2016)), the use of IONM was associated with a significant reduction in postsurgical permanent RILN palsy (Dralle 2004), and RILN monitoring helped to reduce the permanent RILN palsy rate for low-volume surgeons by 0.9% (Sosa 1998). With low-volume surgeons, the permanent RILN palsy rates were highest after visual nerve identification (1.4%) (Sosa 1998).

RILN monitoring might be a useful technique that guides the cautious handling of the recurrent nerve by low-volume surgeons. High-volume surgeons may benefit from RILN monitoring in difficult situations (Dralle 2004). Zheng 2013 published a meta-analysis of 14 different trials, which included 36,487 participants, and concluded that IONM decreases the risk of transient RILN palsy without affecting the rates of permanent injuries.

Description of the intervention

The RILN is normally identified by palpation and surgical dissection. The IONM was introduced in the attempt to identify the nerve by using an electrode (Dequanter 2015). In order to measure the nerve response, the electric field is converted to an acoustic signal, the potentials of which are recorded.

The IONM system operates with two surface electrodes positioned upon an endotracheal tube, which is 7 mm in diameter. During intubation, the anaesthetist inserts, under direct vision, the endotracheal tube between the vocal folds (Figure 1). The RILN is stimulated by a monopolar electrode, using the interrupted stimulation technique (1 mA, 100 ms impulse duration and 4 Hz frequency). In the case of a bifurcated RILN, the post-stimulation response for each nerve branch is included. The endotracheal tube electromyography (EMG) is used to detect the adduction of the vocal folds (Figure 2). A posterior cricoarytenoid muscle contraction, revealed by direct finger palpation, is used to detect the abduction of the vocal folds (Figure 3).

Figure 1. Nerve integrity monitoring endotracheal tube for electromyography signals of a patient's laryngeal muscles (drawn by Silvia Marola)

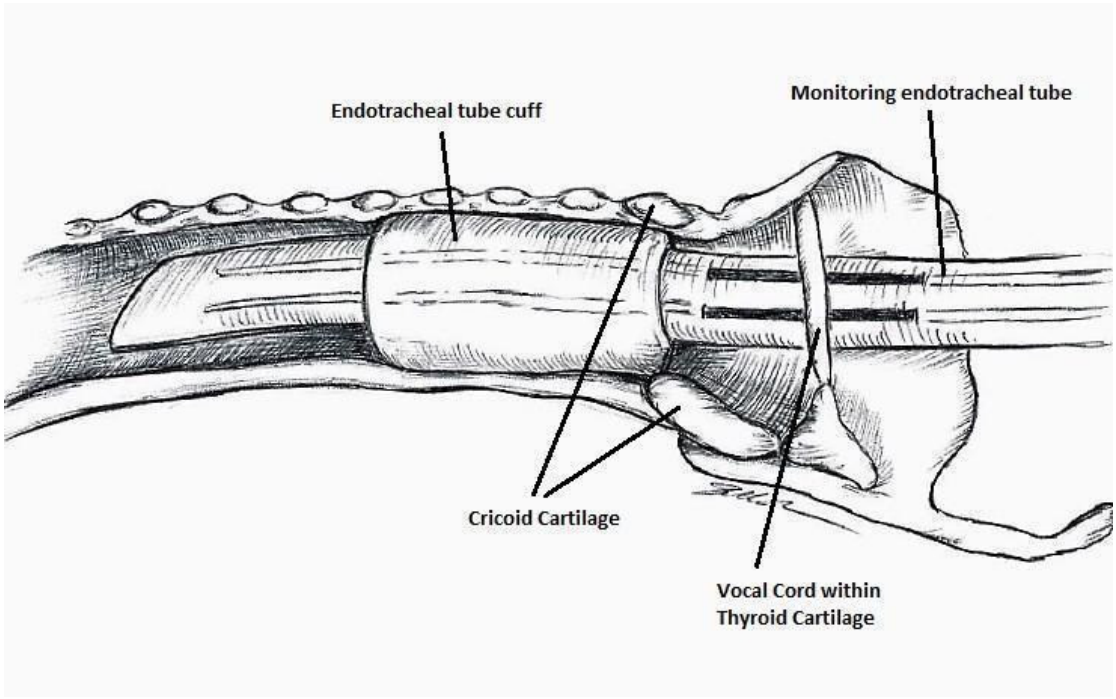


Figure 2. Monitoring endotracheal tube in position positioned at the patient's vocal folds (drawn by Silvia Marola)

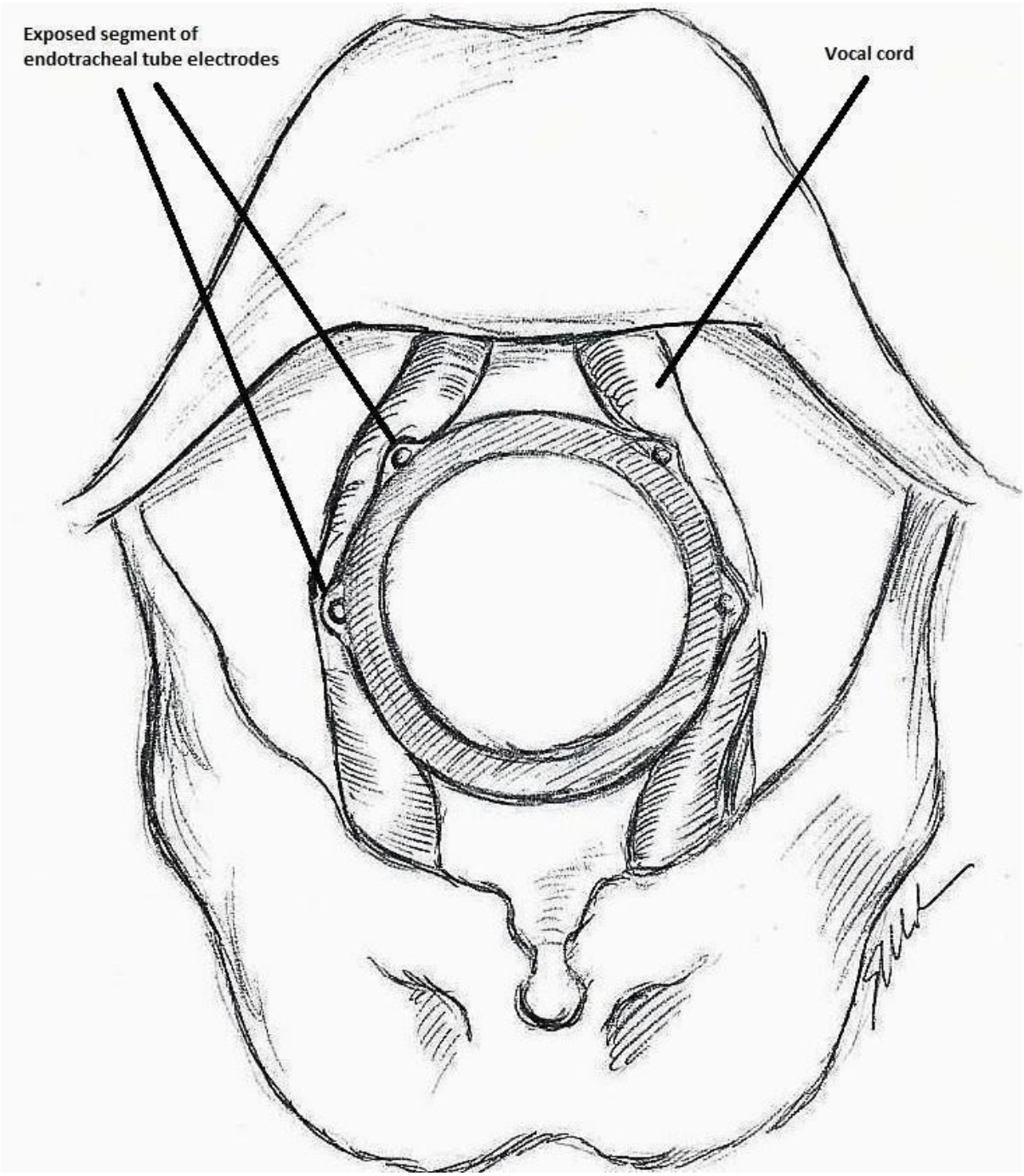
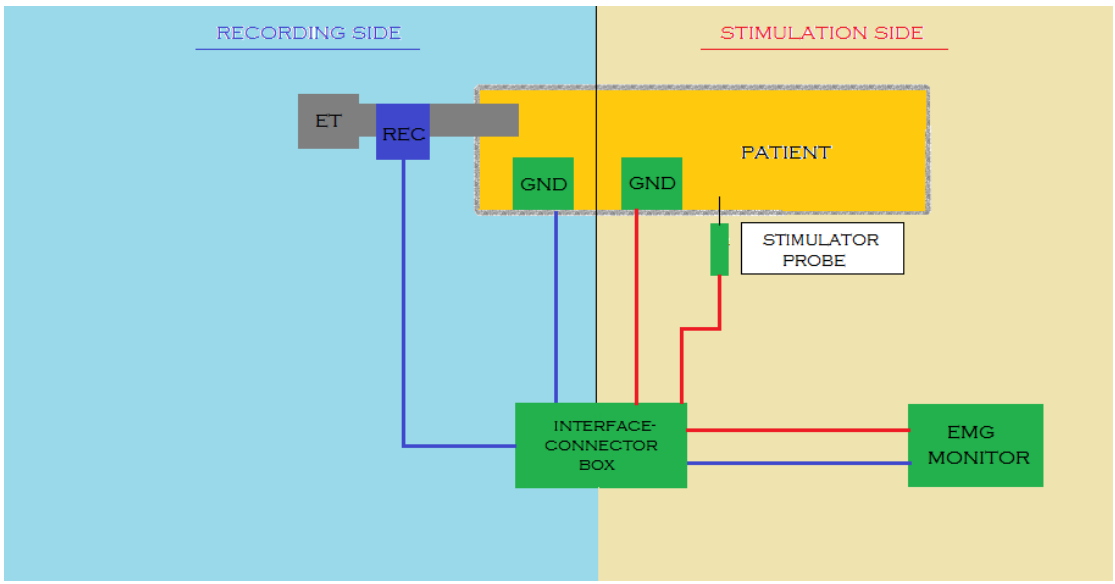


Figure 3. Basic monitoring equipment setup (drawn by Silvia Marola)
EMG: electromyography;ET: endotracheal tube; GND: ground electrodes; REC: recording electrodes



During the intervention, the thyroid lobe is shifted medially and the upper thyroid vessels are tied and cut. The RILN is then identified, dissected and stimulated. At the beginning of thyroidectomy, to ensure that the neuromonitoring system is working, the vagus nerve is stimulated. Proceeding with the operation, the inferior laryngeal nerve is repeatedly stimulated. At the end of the operation, both the vagus and the recurrent nerve are stimulated in an attempt to predict the postoperative outcome (Calò 2014b).
In IONM, the first stimulating electrode is used to make a contact with the RILN, followed by electrical stimulation of the RILN. The second recording electrode receives the electrical signal and a monitor records the signal with a sound. If the RILN is nearby, the surgeon sees the waveform on the displayer at the same time (Zheng 2015). Both the stimulating and recording electrodes are positioned on the sternum or the shoulder, while the interface connector device is linked to the monitoring system. Sonor systems with alarms indicate signal abnormalities, while optic waveform monitors show amplitude, threshold and latency records, which can discriminate true from false signals, giving a real time feedback by monitoring (EMG) responses.
IONM can also be performed in a continuous way, for instance when the stimulating and recording systems are connected to an interface connector device, which is linked to grounding electrodes. Continuous IONM can disclose earlier changes in nerve function, which may be a warning of impending nerve injury (Deniwar 2015a). Continuous IONM seems to be superior to intermittent intraoperative neural monitoring because it enhances standardisation by permanent vagus nerve stimulation, and it provides entire and constant RILN function monitoring as the surgeon dissects and removes the thyroid gland.
Following stimulation of the ipsilateral vagus nerve, the absence of an EMG signal is defined as a loss of signal. An intraoperative algorithm is employed to differentiate between true and false loss of signal. In cases of true loss of signal, the neuromapping technique is used to determine the type of nerve damage and localise the injury site. Following thyroid lobectomy, the loss of signal after vagal stimulation is considered a positive test result. When the laryngoscopy confirms an ipsilateral vocal cord paresis, it is considered a true-positive result. Conversely, a normal mobility of the ipsilateral vocal fold is considered a false-positive result. Following thyroid lobectomy, the detection of a normal signal after vagal stimulation is considered a negative test result. When the postoperative laryngoscopy confirms a normal mobility of the ipsilateral vocal fold, it is considered a true-negative result. Conversely, the detection of an ipsilateral vocal fold paresis is considered as a false-

negative result of the EMG signal.
Co-operation between the surgeon and the anaesthesiologist is essential for successful neuromonitoring. The use of neuromuscular blocking agents should be carefully considered and avoided if possible, as they reduce response amplitudes from the vagus nerve, the RILN and the external branch of the superior laryngeal nerve, which may hinder injury detection.

Adverse effects of the intervention

Chen and colleagues analysed the adverse effects of the procedure of neuromonitoring in a cohort of 3029 patients undergoing thyroid surgery: there were preoperative complications (bucking, deep tracheal catheter placement, tracheal catheter rotation, oversecretion and unstable signal, unstable blood flow dynamics and oral mucosa injury) and postoperative complications (throat pain, pharyngeal discomfort, hoarse voice and joint half-dislocation, inhalation pneumonia, dry eye syndrome, ear and neck numbness and conjunctival congestion (Chen 2015)).

How the intervention might work

During surgery for thyroid carcinoma with lymph node dissection, thyroid reoperation surgery, or in the presence of anatomic variability, IONM can help surgeons to identify the RILN (Dequanter 2015), and may offer a real benefit for lowering nerve injury rates (Malik 2016). An intact monitoring signal at the end of the surgery is associated with a positive outcome for vocal cord functionality. The negative predictive value of the procedure is very high (97% to 99%; Calò 2014b). This means that if 100 patients have an intact monitoring signal at the end of the surgery, 97 to 99 patients out of these 100 patients will have normal vocal cord functionality. On the other hand, with a loss of signal at the end of the operation, the positive predictive value of the procedure is low (33% to 37%), and the occurrence of vocal cord palsy is unpredictable (Calò 2014b). This means that if 100 patients have a loss of signal at the end of the operation, 33 to 37 out of these 100 patients will have vocal cord palsy. During thyroid surgery, neuromonitoring facilitates the identification of RILN, verifying the functional integrity of the nerve (Chiang 2011; Dequanter 2015). In fact, a positive IONM can demonstrate intact nerve function intraoperatively (Chiang 2010). Because most of the injured nerves appear intact, IONM can properly prognosticate postoperative nerve function, which is difficult to detect by visual identification (Dralle 2004). Neuromonitoring can also detect anatomical variation and abnormal courses of the nerves, which are at high risk of injury if not detected (Deniwar 2015b).

Why it is important to do this review

Recent guidelines from the American Academy of Otolaryngology Head and Neck Surgery recommend IONM use in thyroid surgery to prevent nerve damage (Chandrasekhar 2013). IONM is currently used in 80% of thyroidectomies performed by neck surgeons and by more than 50% of general surgeons in the USA. It is more commonly used by higher-volume surgeons (Al-Qurayshi 2016). More clinical trials are needed to further clarify the effects of IONM. In the meta-analysis by Higgins 2011a and in the recent meta-analysis by Pisanu 2014, IONM and visual nerve identification did not demonstrate a substantial difference in rates of transient, total or persistent vocal fold palsy. Pisanu 2014 analysed 20 trials that included 23,152 participants and showed that overall RILN palsy rates for IONM versus visualisation alone were 3.5% versus 3.7%. The role of IONM during thyroid surgery is still debatable, as no consensus exists regarding the prevention of recurrent nerve injury (Deniwar 2015a). There are three primary reasons that this review improves upon the previous reviews by Higgins 2011a and Pisanu 2014. First, both Higgins 2011a and Pisanu 2014 highlighted the need for more trials on this topic that have fewer methodological flaws; both reviews called for more, and better-controlled, randomised controlled trials (RCTs). Pisanu 2014 specifically called for further trials that include high-quality, multicentre, prospective, randomised trials based on strict criteria of standardisation and subsequent meta-analysis to verify the outcomes of interest. If those calls have been heeded, we would expect this Cochrane Review to include more trials with better methodological quality. Second, we included trials published since the last search in August 2013 by Pisanu 2014. Finally, we propose to investigate additional patient-important outcomes that have not yet been investigated in previous reviews, such as health-related quality of life, all-cause mortality and socioeconomic effects.

OBJECTIVES

To assess the effects of IONM versus visual nerve identification for prevention of RILN injury in adults undergoing thyroid surgery.

METHODS

Criteria for considering studies for this review

Types of studies

We included randomised controlled trials (RCTs).

Types of participants

Adults (older than 18 years) undergoing thyroidectomy.
A thyroidectomy is an operation that involves the surgical removal of all or part of the thyroid gland. We evaluated two techniques in this review: partial and total thyroidectomy. We defined partial

thyroidectomy as the surgical removal of a portion of the thyroid gland and total thyroidectomy as the surgical removal of the entire gland.

Types of interventions

We planned to investigate the following comparison of intervention versus control/comparator.

- Intervention: IONM with and without visual nerve identification during thyroidectomy
- Comparator: visual nerve identification only during thyroidectomy

Concomitant interventions had to be the same in both the intervention and comparator groups to establish fair comparisons.

Minimum duration of follow-up

Minimal duration of follow-up was six months.
We defined extended follow-up periods (also called 'open-label extension studies') as follow-up of participants once the original trial, as specified in the trial protocol, had been terminated. However, such trials are frequently of an observational nature and we only planned to evaluate them for adverse events (Buch 2011; Megan 2012).

Specific exclusion criteria

- Clinical trials evaluating people with a previous history of neck surgery and laryngeal nerve injury.

Types of outcome measures

We did not exclude a trial only on the basis that one or more of our primary or secondary outcome measures were not reported in the publication. In the case that none of our primary or secondary outcomes were reported in the trial, we did not include the trial in the synthesis and planned to provide some basic information in an additional table.

Primary outcomes

- Permanent RILN palsy
- Transient RILN palsy
- Health-related quality of life

Secondary outcomes

- Adverse events other than permanent or transient RILN palsy
- Operative time
- All-cause mortality
- Socioeconomic effects

Method of outcome measurement

For nerve-related outcomes (i.e. transient and permanent nerve palsy), the nerve was the unit of analysis. For other outcomes, the participant was the unit of analysis.

- Permanent RILN palsy: defined as an injury detected clinically, by laryngoscopy or both, evaluating the motility of the vocal cords.
- Transient RILN palsy: defined as an injury detected clinically, by laryngoscopy or both, evaluating the motility of the vocal cords.
- Health-related quality of life: evaluated by a validated instrument such as the Short Form 36 (SF-36) or Quality of Life-Thyroid Version (QOL-TV).
- Adverse events: defined as procedure-related events other than permanent or transient RILN palsy, such as hypoparathyroidism. We did not define events not related to the procedure as 'adverse events'. For example, we did not define infection and seroma as adverse events here because these are associated with the closure of skin (i.e. different types of wound closures such as staples or other sutures and comorbidities such as diabetes, corticosteroid treatment, obesity) and not the technique of RILN identification.
- Operative time: defined as the time from the first skin incision to skin closure.
- All-cause mortality: defined as death from any cause.
- Socioeconomic effects: defined as direct costs, including those related to surgical supplies and to hospital stay.

Timing of outcome measurement

The criteria for the timing of outcome measures are listed below.

- Permanent and transient RILN palsy: within six months after surgery
- Health-related quality of life and socioeconomic effects: at 30 days and thereafter
- Adverse events other than permanent or transient RILN palsy: measured at any time after participants were randomised to intervention/comparator groups
- Operative time: at the end of the operation
- All-cause mortality: during the first 30 days after the operation (early mortality) or after 30 days (late mortality)

Search methods for identification of studies

Electronic searches

We searched the following sources from inception of each database to the specified date and placed no restrictions on the language of publication:

- Cochrane Central Register of Controlled Trials (CENTRAL; 2018, Issue 8) via the Cochrane Register of Studies Online (CRSO) (searched 21 August 2018)
- MEDLINE Ovid (Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R); from 1946 onwards; searched 21 August 2018).
- Embase Ovid (from 1974 onwards; searched 12 January 2017). RCTs indexed in Embase are now prospectively added to CENTRAL via a highly sensitive screening process (CENTRAL creation details).
- ClinicalTrials.gov (www.clinicaltrials.gov; searched 21 August 2018).
- World Health Organization International Clinical Trials Registry Platform (ICTRP; www.who.int/trialsearch/; searched 21 August 2018).

Searching other resources

We attempted to identify other potentially eligible trials or ancillary publications by searching the reference lists of included trials, systematic reviews, meta-analyses and health technology assessment reports. In addition we contacted the authors of included trials to identify any additional information on the retrieved trials and to determine if further trials existed that we may have missed. We did not use abstracts or conference proceedings for data extraction unless full data were available from trial authors. Our rationale for this was because this type of information source does not fulfil the CONSORT's "evidence-based, minimum set of recommendations for reporting randomized trials" (CONSORT 2016; Scherer 2007). Rather, we planned to present information on abstracts or conference proceedings in a Characteristics of studies awaiting classification table.

Data collection and analysis

Selection of studies

Two review authors (RC and MB) independently screened the abstract or title, or both, of every record retrieved to determine which trials we should assess further. We investigated the full- text articles of all potentially relevant articles. We resolved discrepancies through consensus or by adjudication by a third review author (VD). If we could not resolve a disagreement, we categorised the trial as a 'study awaiting classification' and contacted the trial authors for clarification. We presented an adapted PRISMA flow diagram to show the process of trial selection (Liberati 2009). We listed all articles excluded after full-text assessment in the Characteristics of excluded studies table and provided the reasons for exclusion.

Data extraction and management

For trials that fulfilled inclusion criteria, two review authors (RC and VD) independently extracted participant and intervention characteristics. We reported data on efficacy outcomes and adverse events using standard data extraction sheets from Cochrane Metabolic and Endocrine Disorders (CMED). We resolved any disagreements by discussion or, if required, we consulted a third review author (AA). For details see Characteristics of included studies; Table 1; Appendix 2; Appendix 3; Appendix 4; Appendix 5; Appendix 6; Appendix 7; Appendix 8; Appendix 9; Appendix 10; Appendix 11; Appendix 12; Appendix 13). GP extracted and evaluated all trial information relating to Barczynski 2009 and Barczynski 2012. The Co-ordinating Editor of the CMED Group checked all data extractions and data analyses.

We provided information about potentially relevant ongoing trials, including trial identifiers, in the Characteristics of ongoing studies table and in Appendix 7, 'Matrix of trial endpoint (publications and trial documents)'. We tried to obtain the protocol for each included trial and reported primary, secondary and other outcomes in comparison with data in the publications listed in Appendix 7. We emailed all authors of included trials to inquire whether they were willing to answer questions regarding their trials. We presented the results of this survey in Appendix 14. We thereafter sought relevant missing information on the trial from the primary trial author(s), if required.

Dealing with duplicate and companion publications

In the event of duplicate publications, companion documents or multiple reports of a primary trial, we maximised the information yield by collating all available data and we used the most complete dataset aggregated across all known publications. We listed duplicate publications, companion documents, multiple reports of a primary trial and trial documents of included trials (such as trial registry information), as secondary references under the study identifier (ID) of the included trial. Furthermore, we also listed duplicate publications, companion documents, multiple reports of a trial and trial documents of excluded trials (such as trial registry information), as secondary references under the study ID of the excluded trial.

Data from clinical trials registers

If data from included trials were available as trial results in clinical trials registers, such as ClinicalTrials.gov or similar sources, we made full use of this information and extracted the data. If there was also a full publication of the trial, we collated and critically appraised all available data. If an included trial was marked as a completed trial in a clinical trials register but no additional information (e.g. trial results, a publication or both) was available, we added this trial to the Characteristics of studies awaiting classification table.

Assessment of risk of bias in included studies

Two review authors (RC and VD) independently assessed the risk of bias of each included trial. We resolved any disagreements by consensus, or by consulting a third review author (AA). In cases of disagreement, we consulted the rest of the review author team and made a judgement based on consensus. If adequate information was unavailable from the trials, trial protocols or both, we contacted the trial authors to recover missing data on 'Risk of bias' items.

We used the Cochrane 'Risk of bias' assessment tool (Higgins 2011b; Higgins 2017), assigning assessments of low, high or unclear risk of bias to each trial (for details see Appendix 2; Appendix 15). We evaluated individual bias items as described in the *Cochrane Handbook for Systematic Reviews of Interventions* according to the criteria and associated categorisations contained therein (Higgins 2017).

Summary assessment of risk of bias

We presented a 'Risk of bias' graph and a 'Risk of bias' summary figure.

We distinguished between self-reported and investigator-assessed and adjudicated outcome measures.

We considered the following to be self-reported outcomes.

- Permanent RILN palsy
- Transient RILN palsy
- Health-related quality of life
- Adverse events other than permanent or transient RILN palsy

We considered the following outcomes to be investigator-assessed.

- Permanent RILN palsy
- Transient RILN palsy
- Adverse events other than permanent or transient RILN palsy
- All-cause mortality
- Operative time
- Socioeconomic effects

Risk of bias for a trial across outcomes

Some 'Risk of bias' domains, such as selection bias (sequence generation and allocation sequence concealment), affect the risk of bias across all outcome measures in a trial. In the case of high risk for selection bias, we marked all endpoints investigated in the associated trial as 'high risk'. Otherwise, we did not perform a summary assessment of the risk of bias across all outcomes for a trial.

Risk of bias for an outcome within a trial and across domains

We assessed the risk of bias for an outcome measure by including all entries relevant to that outcome (i.e. both trial-level entries and outcome-specific entries). We used the term 'low risk of bias' to denote a low risk of bias for all domains, 'unclear risk', to denote an unclear risk of bias for one or more domains, and 'high risk', to denote a high risk of bias for one or more domains.

Risk of bias for an outcome across trials and across domains

This type of risk of bias was the type of main summary assessments that we incorporated into our judgements about the quality of evidence in the 'Summary of finding' table. We defined outcomes as being at 'low risk of bias' when most information came from trials at low risk of bias, 'unclear risk', when most information came from trials at low or unclear risk of bias, and 'high risk', when most information came from trials at high risk of bias.

Measures of treatment effect

When at least two included trials were available for a comparison on a given outcome, we expressed dichotomous data as a risk ratio (RR) with 95% confidence intervals (CIs). For continuous outcomes measured on the same scale (e.g. surgical time), we estimated the intervention effect using the mean difference (MD) with 95% CIs. For continuous outcomes that measured the same underlying concept (e.g. health-related quality of life), but used different measurement scales, we planned to calculate standardised mean difference (SMD). We planned to express time-to-event data as a hazard ratio (HR) with 95% CIs.

Unit of analysis issues

We took into consideration the level at which randomisation occurred, such as cross-over trials, cluster-randomised trials and multiple observations for the same outcome. If more than one comparison from the same trial was eligible for inclusion in the same meta-analysis, we either planned to combine groups to create a single pair-wise comparison, or appropriately reduced the sample size so that the same participants did not contribute multiply (e.g. splitting the 'shared' group into two or more groups). While the latter approach offers some solution to adjusting the precision of the comparison, it does not account for correlation arising from the same set of participants being in multiple comparisons (Higgins 2011c).

If applicable, we planned to reanalyse cluster-RCTs that did not appropriately adjust for potential clustering of participants within clusters in their analyses. We intended to inflate the variance of the intervention effects using design effects such as an estimation of an intra-cluster correlation (ICC). Where applicable, we planned to obtain estimates of ICCs through contact with the trial authors, or imputed them using estimates from other included trials that reported ICCs, or using external estimates from empirical research (e.g. Bell 2013). We planned to examine the impact of clustering using sensitivity analyses if necessary.

Dealing with missing data

If possible, we obtained missing data from the authors of the included trials. We carefully evaluated important quantitative data such as the number of screened, randomly assigned participants, as well as intention-to-treat, as-treated and per-protocol populations. We investigated attrition rates (e.g. dropouts, losses to follow-up, withdrawals), and we critically appraised issues concerning missing data and use of imputation methods (e.g. last observation carried forward).

Where included trials reported median and ranges instead of means and standard deviations (SD) for outcomes, and we did not receive the necessary information from trial authors, we imputed these values by estimating the mean and variance from the median, range and the size of the sample when trial authors reported those nonparametric statistics (Hozo 2005). In trials where the SD of the outcome was not available at follow-up but where other trials reported this information, we standardised, by the average of the pooled baseline SD, from those trials that reported this information.

We investigated the impact of imputation on meta-analyses by performing sensitivity analyses and we reported for each outcome which trials were included with imputed SDs.

Assessment of heterogeneity

In the event of substantial clinical or methodological heterogeneity, we reported trial results in the pooled effect estimate.

We identified heterogeneity (inconsistency) by visually inspecting forest plots and by using a standard Chi² test with a significance level of $\alpha = 0.1$ (Deeks 2017). In view of the low power of this test, we also considered the I^2 statistic (Higgins 2003), which quantifies inconsistency across trials to assess the impact of heterogeneity on the meta-analysis (Higgins 2002). An I^2 statistic value greater than 75% indicates a considerable level of heterogeneity (Deeks 2017). When we found heterogeneity, we attempted to determine the possible reasons for it by examining individual trial and subgroup characteristics.

Assessment of reporting biases

If we included 10 or more trials that investigated a particular outcome, we used funnel plots to assess small-trial effects. Several explanations may account for funnel plot asymmetry, including true heterogeneity of effect with respect to trial size, poor methodological design (and hence bias of small trials) and publication bias. Therefore, we interpreted the results carefully (Sterne 2011).

Data synthesis

We planned to undertake (or display) a meta-analysis only if we judged participants, interventions, comparisons, and outcomes to be sufficiently similar to ensure that the results had a clinically meaningful interpretation. Unless sufficient evidence showed homogeneous effects across trials of different methodological quality, we primarily summarised low risk of bias data using a random-effects model (Wood 2008). We interpreted random-effects meta-analyses with due consideration to the whole distribution of effects and presented a prediction interval (Borenstein 2017a; Borenstein 2017b; Higgins 2009). A prediction interval needs at least three trials to be calculated and specifies a predicted range for the true treatment effect in an individual trial (Riley 2011). For rare events, such as event rates below 1%, we planned to use the Peto's odds ratio method, provided that there was no substantial imbalance between intervention and comparator group sizes and intervention effects were not exceptionally large. In addition, we performed statistical analyses according to the statistical guidelines presented in the *Cochrane Handbook for Systematic Reviews of Interventions* (Deeks 2017).

Subgroup analysis and investigation of heterogeneity

We expected the following characteristics to introduce clinical heterogeneity and planned to carry out the following subgroup analyses including investigation of interactions when applicable (Altman 2003).

- Partial versus total thyroidectomy
- Thyroidectomy for cancer versus benign thyroid disease
- Low versus high experience in thyroid surgery; we defined 'low experience in thyroid surgery' as case-volume of thyroidectomies less than 25 per year (Adam 2016).
- Residents in general surgery versus surgeons
- Participants aged less than 75 years versus 75 years or older
- Participants with a body mass index of less than 35 kg/m² versus body mass index 35 kg/m² or higher
- Thyroidectomy with tie and clamp versus vascular dissection, cutting and sealing simultaneously (UltraCision Harmonic® scalpel) or a bipolar vascular sealing system (LigaSure®)

Sensitivity analysis

We planned to perform sensitivity analyses to explore the influence of the following factors (when applicable) on effect sizes based on the following characteristics.

- The publication status of the trial
- The effect of risk of bias, as specified in the Assessment of risk of bias in included studies section
- The extent to which very large trials dominated the results
- And other criteria such as diagnostic criteria, imputation, language of publication, source of funding (industry versus other), or country

We also tested the robustness of results by repeating the analyses using different measures of effect size (RR, OR, etc), and different statistical models (fixed-effect and random-effects models).

Certainty of evidence

We presented the overall certainty of the evidence for each outcome specified in the Types of outcome measures section according to the GRADE approach, which takes into account issues related not only to internal validity (risk of bias, inconsistency, imprecision, publication bias) but also to external validity, such as the directness of results. Two review authors (IA and RC) independently rated the certainty of evidence for each outcome. We resolved differences in assessment by discussion or consultation with a third review author (NA).

We included the 'Checklist to aid consistency and reproducibility of GRADE assessments' (Appendix 15), to help with standardisation of the 'Summary of findings' tables (Meader 2014). Alternatively, we planned to use the GRADEpro Guideline Development Tool (GDT) software and would have presented evidence profile tables as an appendix (GRADEproGDT 2015). We presented the results for the outcomes as described in the Types of outcome measures section. If meta-analysis was not possible, we presented the results in a narrative format in the 'Summary of findings' table. We justified all decisions to downgrade the certainty of the evidence using footnotes and we made comments to aid the reader's understanding of the Cochrane Review where necessary.

'Summary of findings' table

We presented a summary of the evidence in Summary of findings for the main comparison. This information provides key information about the best estimate of the magnitude of the effect, in relative terms and as absolute differences, for each relevant comparison of alternative management strategies, numbers of participants and trials addressing each important outcome, and a rating of overall confidence in effect estimates for each outcome. We created the 'Summary of findings' table based on the methods described in the *Cochrane Handbook for Systematic Reviews of Interventions* (Schünemann 2017) using Review Manager 5 (RevMan 5) table editor (Review Manager 2014). We reported the following outcomes, listed according to priority.

- Permanent RILN palsy
- Transient RILN palsy
- Health-related quality of life
- Adverse events other than permanent or transient RILN palsy (transient hypoparathyroidism)
- Operative time
- All-cause mortality
- Socioeconomic effects

RESULTS

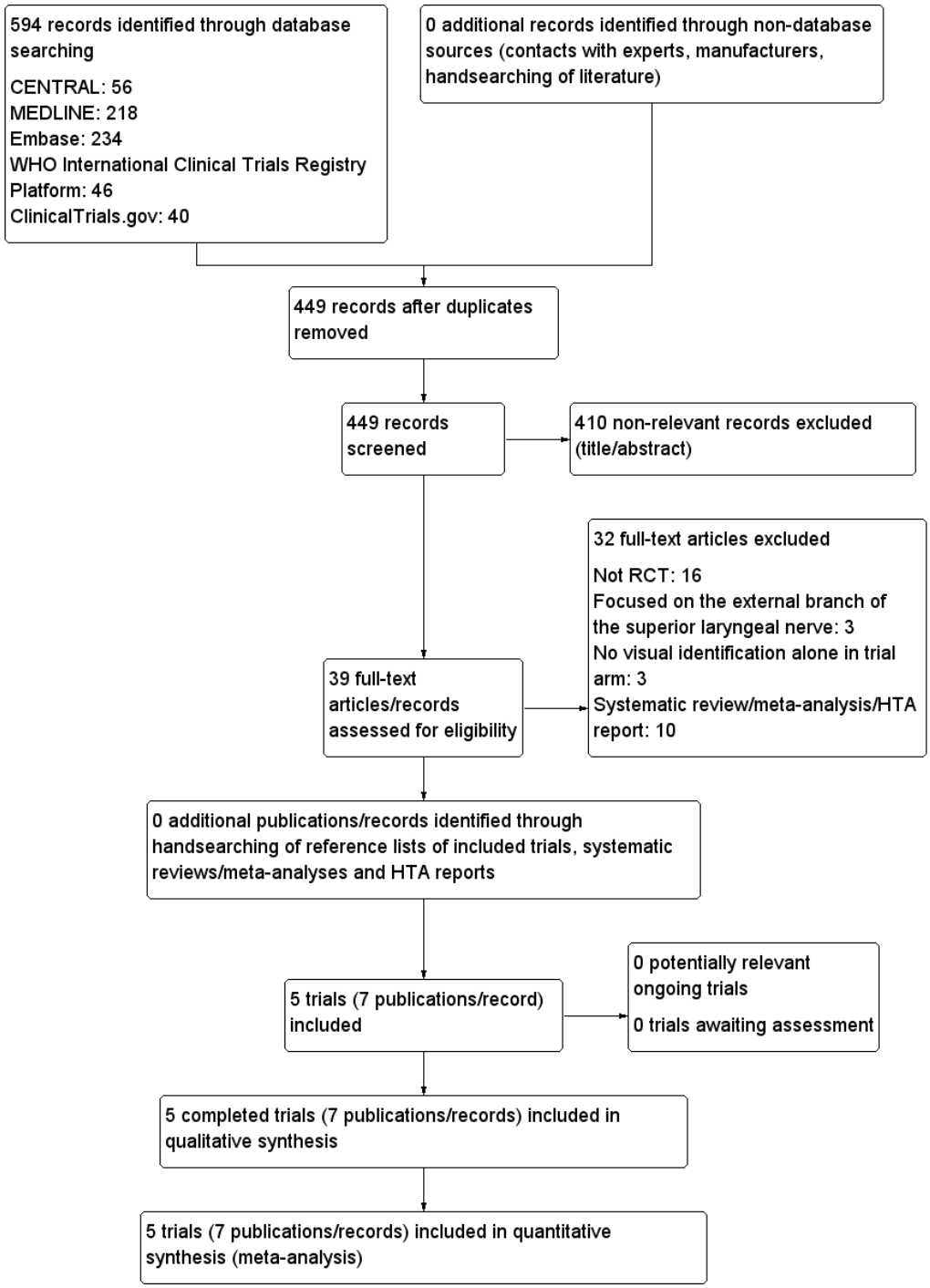
Description of studies

For a detailed description of trials, see Table 1, Characteristics of included studies, Characteristics of excluded studies, and Characteristics of ongoing studies.

Results of the search

In our comprehensive literature searches, we identified a total of 594 records. There were 447 records after the exclusion of duplications. From these, we identified 37 full-text publications for further examination. We excluded the other publications on the basis of titles or abstracts because inclusion criteria had not been met or because the trials were not relevant to the review objectives (see Figure 4 for the amended PRISMA study flow diagram). After screening the full text of the selected publications, five trials (Barczynski 2009; Barczynski 2012; Hei 2016a; Lee 2015; Sari 2010), met our inclusion criteria. All trials were published in English.

Figure 4. Trial flow diagram



Included studies

For a detailed description of trials, see the Table 1 and the Characteristics of included studies section.

Source of data

We identified 594 publications: 56 in CENTRAL, 218 in MEDLINE, 234 in Embase, 46 in the WHO ICTRP and 40 in ClinicalTrials.gov; we did not identify any additional records through handsearching reference lists. After removal of duplicates, we evaluated 447 publications.

Comparisons

All the included trials compared IONM plus visual nerve identification with visual nerve identification alone for prevention of RILN in participants undergoing conventional thyroidectomy (Barczynski 2009; Barczynski 2012; Hei 2016a; Sari 2010), or robotic thyroidectomy (Lee 2015).

Overview of trial populations

The five trials included a total of 1558 participants undergoing thyroidectomy: 781 participants were randomised to IONM with visual nerve identification and 777 to visual nerve identification only. Trial sample size ranged from 50 (Lee 2015), to 1000 (Barczynski 2009).

Trial design

All five trials utilised a parallel-group, superiority design. All trials were monocentric. One of the trials (Barczynski 2012), was double-blinded (participants and the ear, nose and throat specialist who was performing videostrobolaryngoscopy and voice assessment). The trials were performed between the years 2006 and 2014. The mean duration of follow-up ranged from 6 (Barczynski 2012; Hei 2016a), to 12 months (Barczynski 2009; Lee 2015; Sari 2010). No trial was terminated early.

Settings

Two trials were conducted in Poland (Barczynski 2009; Barczynski 2012) and one trial each was conducted in China (Hei 2016a), Korea (Lee 2015), and Turkey (Sari 2010). All trials were conducted at academic institutions and had an inpatient setting.

Participants

Participants had an indication for thyroidectomy due to a variety of diseases including toxic and nontoxic nodular goitre, Graves' disease and thyroid carcinoma (Appendix 5). Three hundred-seven (18.7%) participants came from low- and middle-income countries such as China (Hei 2016a) and Turkey (Sari 2010). In three trials the ethnic groups were white participants (Barczynski 2009; Barczynski 2012; Sari 2010), in two trials there were Asian participants (Hei 2016a; Lee 2015). None of the trials reported any information on the duration of thyroid disease. Surgical procedures were most commonly performed with female participants (81%). In the IONM group the mean age of trial participants ranged from 44.2 years (Lee 2015), to 51.3 years (Barczynski 2009). In the visual identification only group the mean age of trial participants ranged from 41.7 years (Lee 2015), to 51.9 years (Barczynski 2009; Appendix 6). The mean BMI reported in one trial (Sari 2010), was 26.9 kg/m^2 (SD 3) in the IONM group versus 27.3 kg/m^2 (SD 3) in the visual identification only group. No trial reported on comorbidities. Major exclusion criteria differed amongst trials: in Barczynski 2009, Barczynski 2012 and Lee 2015, previous thyroid or parathyroid surgery was an exclusion criterion. By contrast, in Hei 2016a, this was a specific inclusion criterion. In Barczynski 2009, Barczynski 2012 and Hei 2016a participants underwent central neck compartment dissection or lateral neck dissection; their major exclusion criterion was Graves' disease.

Diagnosis

No trial reported diagnostic procedures.

Interventions

No trial reported treatment before the start of the trial.
All five trials used IONM as their intervention. The system used for neuromonitoring of the IONM group was the NIM 2.0/3.0 system® (Medtronic Xomed Surgical Products, Jacksonville, FL), in four trials (Barczynski 2012; Hei 2016a; Lee 2015; Sari 2010), and the Neurosign® 100 system (Inomed, Teningen, Germany), in one trial (Barczynski 2009). In four trials an anaesthetist inserted endotracheal tube surface electrodes between the vocal folds under direct vision during intubation (Barczynski 2009; Barczynski 2012; Lee 2015; Hei 2016b); in the other trial, the trial authors used needle electrodes (Sari 2010). The trials located, mapped, and stimulated RILN using a probe, and confirmed the identification of a healthy RILN by acoustic evaluation of the signal and visual display of the EMG response evaluation (latency and amplitude). All five trials used RILN visualisation only as the comparator intervention.

Outcomes

In Barczynski 2009 and Barczynski 2012 there were no differences between stated primary outcomes in the publication and ClinicalTrials.gov (NCT00661024; NCT01395134). No protocol or trials register information was available for the other three trials.
Four trials (Barczynski 2009; Barczynski 2012; Hei 2016a; Lee 2015), explicitly stated a primary endpoint in the publication. Only Barczynski 2009 and Barczynski 2012 reported secondary endpoints. The defined primary outcomes were the incidence of the RILN injury (evaluated on the second postoperative day and at 1, 2, 4, 6 and 12 months postoperatively, if paresis was noted on first examination; Barczynski 2009), the identification rate of the external branch of the superior laryngeal nerve (evaluated up to six months postoperatively; Barczynski 2012), postoperative RILN function (evaluated by laryngoscopy (Hei 2016a); or transient or permanent unilateral or bilateral laryngeal nerve lesions, voice handicap index and voice range profile (Lee 2015)).
The five trials reported a mean of seven (range 2 to 14) outcomes. Two trials (Lee 2015; Sari 2010), reported the rate of adverse events related to IONM. One trial (Sari 2010), reported all-cause mortality; Barczynski 2009 and Barczynski 2012 authors confirmed that there were no deaths up to 30 days after surgery in their trials. No trial assessed health-related quality of life or socioeconomic effects.

Excluded studies

After careful evaluation of the full publication we had to exclude 32 trials. The main reasons for exclusion were that the trial design was not an RCT and there was no adequate intervention or control group (for details see Characteristics of excluded studies).

Risk of bias in included studies

For details on the risk of bias of the included trials see Characteristics of included studies, Appendix 2 and Appendix 3. For an overview of review authors' judgements about each risk of bias item for individual trials and across all trials see Figure 5 and Figure 6.

Figure 5. 'Risk of bias' summary: review authors' judgements about each risk of bias item for each included trial ((blank cells indicate that the particular outcome was not measured in the associated trial)

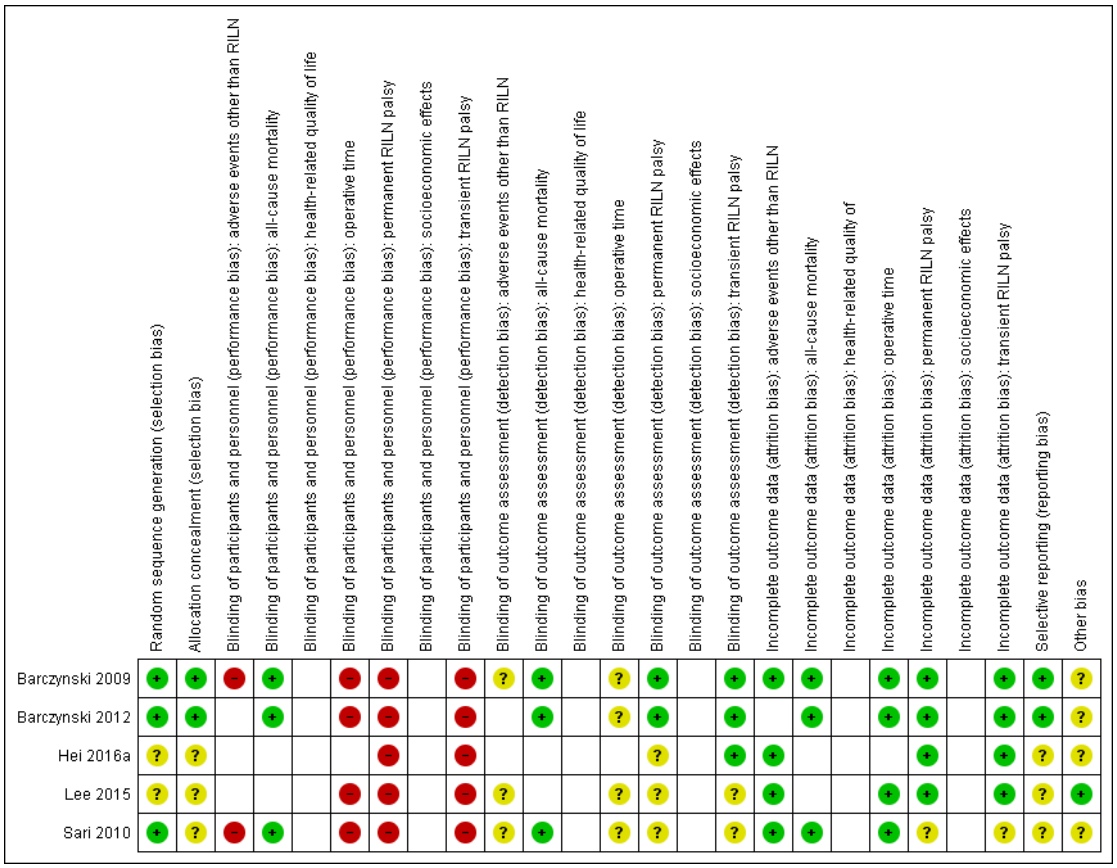


Figure 6. 'Risk of bias' graph: review authors' judgements about each risk of bias item presented as percentages across all included trials (blank cells indicate that the particular outcome was not measured in some or all trials)

Allocation

Random sequence generation was well explained in three trials (Barczynski 2009; Barczynski 2012; Sari 2010), so we judged them at low risk of bias, with the exception of Hei 2016a and Lee 2015 with an unclear process of randomisation.

Two out of five trials reported details of allocation concealment and we considered them as low risk of bias (Barczynski 2009; Barczynski 2012). Three trials had an unclear reporting of the allocation concealment (Hei 2016a; Lee 2015; Sari 2010).

Blinding

No publication provided information about blinding of participants and personnel but we believe personnel were probably not blinded in any trial. We therefore assigned a high risk of performance bias to the trials that reported the outcomes permanent, transient RILN palsy and operative time. In Barczynski 2009, the authors stated that an ear, nose and throat (ENT) specialist who performed laryngeal examinations on postoperative day two and followed up participants in case of nerve injury was blinded to trial conditions. In Barczynski 2012 the authors reported the procedure of the blinding of outcome assessment for adverse events other than RILN ("The assessment protocol was strictly followed by the ENTspecialist (AH), who was blinded to the patient relevant group assignment").

Incomplete outcome data

All included trials specified participants' withdrawals or reasons for withdrawal. They analysed data mostly on an intention-to-treat basis. Sari 2010 excluded 4.3% of enrolled participants in the IONM group due to lack of signal. Barczynski 2012 lost a total of nine participants (4.3%) during follow up without providing reasons in the publication, however, the number of participants lost to follow-up was balanced among treatment arms; therefore, it is unlikely that the low attrition rate affected outcome measures.

Selective reporting

Four trials provided a clinical trial identifier or a reference to a protocol (Barczynski 2009; Barczynski 2012; Hei 2016a; Lee 2015). We could not identify the protocol for Hei 2016a and Lee 2015. The primary outcomes in ClinicalTrials.gov (NCT00661024; NCT01395134), and in the published articles (Barczynski 2009; Barczynski 2012), were the same. The remaining trial (Sari 2010), had an unclear risk of reporting bias due to no protocol being available.

Other potential sources of bias

Four trials measured clustered data for the outcomes transient and permanent RILN palsy but did not use adequate statistical analyses for these data (Barczynski 2009; Barczynski 2012; Hei 2016a; Sari 2010). We, therefore, judged them to be at unclear risk of other bias.

Effects of interventions

See: **Summary of findings for the main comparison** Intraoperative neuromonitoring compared to visual nerve identification only during thyroidectomy

Baseline characteristics

For details of baseline characteristics, see Appendix 5 and Appendix 6.

Intraoperative neuromonitoring with and without visual nerve identification versus visual nerve identification only

Primary outcomes

Permanent RILN palsy

Permanent RILN palsy was a rare event. Comparing IONM with visual nerve identification only for nerves at risk showed a risk ratio (RR) of 0.77 (95% CI 0.33 to 1.77; P = 0.54; 4 trials; 2895 nerves at risk; very low-certainty evidence; Analysis 1.1). Ten of 1451 operated nerves (0.7%) in the IONM group compared with 13 of 1444 operated nerves (0.9%) in the visual identification only group showed permanent nerve damage. For clustered data in trials where the unit of analysis was the nerve and not the participant, we recognised that the resulting 95% CIs produced were probably narrower, and the actual variance greater, than reported here. No trial performed an adequate statistical analysis with regard to the cluster design. Therefore, results of Analysis 1.1 have to be interpreted as an effect estimate with probably even larger 95% CIs than indicated. There were no substantial differences in fixed- versus random-effects RRs. The 95% prediction interval ranged between 0.12 and 4.79.

Transient RILN palsy

IONM compared with visual nerve identification only for nerves at risk, showed a RR of 0.62 (95% CI 0.35 to 1.08; P = 0.09; 4 trials; 2895 nerves at risk; very low-certainty evidence; Analysis 1.2). Thirty of 1451 operated nerves (2.1%) in the IONM group compared with 52 of 1444 operated nerves (3.6%) in the visual identification only group showed transient nerve damage. The use of a fixed-effect model resulted in a RR slightly more in favour of IONM (RR 0.58, 95% CI 0.37 to 0.90). For clustered data in trials where the unit of analysis was the nerve and not the participant, we recognised that the resulting 95% CIs produced were probably narrower, and the actual variance greater, than reported here. None of the included trials performed an adequate statistical analysis with regard to the cluster design. Therefore, results of Analysis 1.2 have to be interpreted as an effect estimate with probably even larger 95% CIs than indicated. The 95% prediction interval ranged between 0.12 and 3.11.

Health-related quality of life

Health-related quality of life was not reported in any trial

Secondary outcomes

Adverse events other than permanent or transient RILN palsy (transient hypoparathyroidism)

Adverse events other than permanent or transient RILN palsy for IONM compared with visual nerve identification only showed a RR of 1.25 (95% CI 0.45 to 3.47; P = 0.66; 2 trials; 268 participants; very low-certainty evidence; Analysis 1.3). Twenty of 147 participants (13.6%) in the IONM group compared with 17 of 139 participants (12.2%) in the visual identification only group experienced an adverse event. Use of a fixed-effect model did not substantially change the effect size.

Operative time

The mean difference (MD) in operative time between the IONM group and the visual identification only group was −0.8 minutes (95% CI −11.2 to 9.6; P = 0.88; 4 trials; 1488 participants; very low-certainty evidence; Analysis 1.4). Use of a fixed-effect model showed a MD of 4.8 minutes (95% CI 2.3 to 7.3), in favour of visual nerve identification only. In Sari 2010, the surgical approach in around 80% of the participants was a lobectomy only. When removing this trial from the meta-analysis, the MD was 5.5 minutes (95% CI −0.7 to 11.8; P = 0.08, 3 trials; 1251 participants). The 95% prediction interval ranged between −60.6 minutes and 71.7 minutes. The fixed-effect model showed a MD of 7 minutes (95% CI 4.3 to 9.6), in favour of visual nerve identification only.

All-cause mortality

In the three trials with 1438 participants providing information on this outcome (Barczynski 2009; Barczynski 2012; Sari 2010), no postoperative hospital deaths or deaths within the first 30 days occurred after thyroid surgery (moderate certainty evidence; Analysis 1.5).

Socioeconomic effects

None of the included trials reported socioeconomic effects.

Subgroup analyses

We did not perform subgroup analyses because there were not enough trials to estimate effects in various subgroups.

Sensitivity analyses

With the exception of random-effects model versus fixed-effect model, we did not perform sensitivity analyses because there were not enough trials. Model choice affected the direction of results for transient RILN palsy and operative time, but the practical differences were minor, nonetheless.

Assessment of reporting bias

We did not draw funnel plots due to the limited number of trials (N = 5).

Ongoing trials

We did not identify any ongoing RCTs.

DISCUSSION

Summary of main results

We did not find compelling evidence for the superiority or inferiority of IONM over visual nerve identification alone on any of the outcomes measured.

Overall completeness and applicability of evidence

While there were five trials that met our inclusion criteria, at most, we could synthesise four trials for any one outcome. Many outcomes were low-occurence events (e.g. all-cause mortality), and there were no data on health-related quality of life and socioeconomic events. There was marked clinical heterogeneity indicating difficult applicability of the results across a variety of clinical or experimental settings. Finally, there was a unit-of-analysis issue for the trials evaluating RILN palsy outcomes, not accounting for the clustering nature of nerves at the trial level.

Quality of the evidence

This review containing five trials does not provide robust evidence regarding IONM for reducing transient or permanent RILN palsy, other adverse events and operative time compared to visual nerve identification alone. We considered the certainty of the evidence for these outcomes to be very low according to GRADE. The reasons for downgrading included risk of bias (three out of five trials had unclear allocation concealment, three out of five trials had unclear reporting bias, all trials were exposed to performance bias), imprecision of results and inconsistency. The only outcome we

considered as moderate-certainty evidence was all-cause mortality; no events occurred but only three trials reported this outcome.

Potential biases in the review process

Although our review was comprehensive in terms of the search strategy, the possible presence of reporting bias can not be excluded. We applied no restrictions or date limitations to the search and had access to all scientific databases. We contacted the trial authors of the included trials and attempted to gather all information required in order to limit this type of bias. However, only Barczynski (Barczynski 2009; Barczynski 2012), replied to our questions. We did not detect relevant departures from trial protocols during the review process.

For all analyses, we opted for a random-effects analysis due to substantial clinical heterogeneity. For Analysis 1.4, we detected Sari 2010 as an outlier (which would give justification to exclude this trial), however, we decided to leave this trial in our analysis. New clinical trials are needed to clarify the situation.

Marginal decisions around the inclusion or exclusion of trials or use and analysis of data could have had an impact on the findings of the review (e.g. clinical heterogeneity, variation in trial design or delivery of intervention, prioritisation of data from multiple time points, definition of subgroups, alternative definitions of outcomes, use of adjusted as opposed to unadjusted data, outcome surrogacy). Analysis of data on primary outcomes involved consideration of clustered data. RILN palsy nerves in the original trials were used by trial authors as the denominator without adjustment for the non-independence between nerves. Also, these trials did not employ adequate statistical measures. Because no intra-cluster correlation data were available, we did not reanalyse clustered data for RILN palsy. Furthermore, CIs were consistent with both benefit and harm and would have been even wider had case re-analyses been possible. Regarding the exclusion of trials, we excluded trials investigating injury of the external branch of the superior laryngeal nerve due to a previous decision reported in the protocol (Cirocchi 2016).

Agreements and disagreements with other studies or reviews

Ten previous systematic reviews and meta-analyses of comparative trials on IONM versus visual nerve identification have been published (Bai 2018; Lombardi 2016; Malik 2016; Pisanu 2014; Rulli 2014; Sanabria 2013; Sun 2017; Wong 2017; Yang 2017; Zheng 2013); see Table 2. Two other meta-analyses (Higgins 2011; Pardal-Refoyo 2016), performed a pooled analysis of comparative and non-comparative trials. Two meta-analyses were performed on a specific subgroup: patients who underwent thyroid reoperations (Sun 2017), and high-risk patients (re-operation, thyroidectomy for malignancy, thyrotoxicosis and retrosternal goitre; Wong 2017).

The results of the previous reviews were heterogeneous and conflicting. There was evidence for a decrease of transient and permanent RILN injury (Bai 2018; Zheng 2013), a decrease of transient RILN injury (Wong 2017; Yang 2017), a decrease of permanent RILN injury (Sun 2017), no substantial decrease of permanent RILN injury (Lombardi 2016), and no substantial decrease for transient and permanent RILN injury (Malik 2016; Pisanu 2014; Sanabria 2013).

The reasons for these differences may be explained by the continuous evolution and progress of IONM technology and the rising surgical awareness for the need of strict adherence to a standardised approach to IONM, as outlined in the published guideline statements (Barczynski 2012; Randolph 2011). In the older trials, sensitivity and specificity in RILN identification was probably lower. Moreover, the use of a non-standardised approach to IONM procedures may have limited the potential of the method in prognostication of postoperative neural function. Only a few centres used routine postoperative laryngoscopy, whereas "no hoarseness" reported by the patient or found on clinical examination served as sufficient proof of intact vocal folds and RILN function in the majority of centres. This may have resulted in an underestimation of the incidence of RILN injury (Henry 2017). Moreover, assessment of novel technology of continuous vagal IONM, which allows for recognition of impending neural injury by EMG change and has the potential to prevent neural injury by modification of surgical manoeuvres, was outside the scope of the majority of published systematic reviews and meta-analyses.

AUTHORS' CONCLUSIONS

Implications for practice

Both intraoperative neuromonitoring (IONM) and visual nerve identification alone are widely used during thyroidectomy. The choice of technique depends on various factors, such as technical resources and the surgeon's experience. However, it is unclear if IONM should be preferred to visual nerve identification only in all patients undergoing thyroid surgery or if it should be used rather in selected patients, at high-risk for recurrent inferior laryngeal nerve (RILN) injury (e.g. revision thyroid surgery). As the literature on this topic to date is limited in quantity and quality, there is no robust evidence that supports a substantive difference between interventions in the analysis of permanent and transient RILN palsy, adverse events other than RILN palsy (transient hypoparathyroidism), and operative time. In general, the use of IONM is as safe as visual nerve identification alone. More research is needed to confidently determine whether IONM reduces transient or permanent types of palsy, with a focus on risk stratification for unilateral versus bilateral nerve events.

Implications for research

The major limitations of our meta-analyses are the small sample sizes and the overall low number of events in the included trials. Well-designed, executed, analysed and reported randomised controlled trials, with a larger number of participants and longer follow-up, employing the latest IONM technology and applying new surgical techniques (e.g. staged thyroidectomy in case of loss of signal on the first side (Henry 2017)), are needed.

ACKNOWLEDGEMENTS

The authors thank Maria-Inti Metzendorf, the Cochrane Metabolic and Endocrine Disorders' Information Specialist (IS), for the development of the search strategy.

The authors thank Silvia Marola for drawing figures one to three.

We thank our peer reviewer Bianca Heemingsen for assisting with the preparation of the review.

REFERENCES

References to studies included in this review

Barczynski 2009 *{published data only}*

* Barczynski M, Konturek A, Cichon S. Randomized clinical trial of visualization versus neuromonitoring of recurrent laryngeal nerves during thyroidectomy. *The British Journal of Surgery* 2009;**96**(3):240–6. PUBMED: 19177420]

NCT00661024. Visualization versus intraoperative neuromonitoring of the recurrent laryngeal nerves in thyroid surgery. https://clinicaltrials.gov/ct2/show/NCT00661024 (accessed 20 December 2018).

Barczynski 2012 *{published data only}*

* Barczynski M, Konturek A, Stopa M, Honowska A, Nowak W. Randomized controlled trial of visualization versus neuromonitoring of the external branch of the superior laryngeal nerve during thyroidectomy. *World Journal of Surgery* 2012;**36**(6):1340–7. PUBMED: 22402975]

NCT01395134. Visualization versus neuromonitoring of the external branch of the superior laryngeal nerve during thyroidectomy. https://clinicaltrials.gov/ct2/show/NCT01395134 (accessed 20 December 2018).

Hei 2016a *{published data only}*

Hei H, Zhou B, Qin J, Song Y. Intermittent intraoperative nerve monitoring in thyroid reoperations: preliminary results of a randomized, single-surgeon study. *Head & Neck* 2016;**38 Suppl 1**:E1993–7. PUBMED: 26829763]

Lee 2015 *{published data only}*

Lee HY, Lee JY, Dionigi G, Bae JW, Kim HY. The efficacy of intraoperative neuromonitoring during robotic thyroidectomy: a prospective, randomized case-control evaluation. *Journal of Laparoendoscopic & Advanced Surgical Techniques* 2015;**25**(11):908–14. PUBMED: 26575249]

Sari 2010 *{published data only}*

Sari S, Erbil Y, Sumer A, Agcaoglu O, Bayraktar A, Issever H, et al. Evaluation of recurrent laryngeal nerve monitoring in thyroid surgery. *International Journal of Surgery* 2010;**8**(6):474–8. PUBMED: 20601257]

References to studies excluded from this review

Alesina 2012 *{published data only}*

Alesina PF, Rolfs T, Hommeltenberg S, Hinrichs J, Meier B, Mohmand W, et al. Intraoperative neuromonitoring does not reduce the incidence of recurrent laryngeal nerve palsy in thyroid reoperations: results of a retrospective comparative analysis. *World Journal of Surgery* 2012;**36**(6):1348–53. PUBMED: 22411090]

Barczynski 2011 *{published data only}*

Barczynski M, Konturek A, Stopa M, Hubalewska-Dydejczyk A, Richter P, Nowak W. Clinical value of intraoperative neuromonitoring of the recurrent laryngeal nerves in improving outcomes of surgery for well-differentiated thyroid cancer. *Polski Przeglad Chirurgiczny* 2011;**83**(4):196–203. PUBMED: 22166358]

Brauckhoff 2002 *{published data only}*

Brauckhoff M, Gimm O, Thanh PN, Brauckhoff K, Ukkat J, Thomusch O, et al. First experiences in intraoperative neurostimulation of the recurrent laryngeal nerve during thyroid surgery of children and adolescents. *Journal of Pediatric Surgery* 2002;**37**(10):1414–8. PUBMED: 12378445]

Calo 2013 *{published data only}*

Calo PG, Pisano G, Medas F, Tatti A, Pittau MR, Demontis R, et al. Intraoperative recurrent laryngeal nerve monitoring in thyroid surgery: is it really useful?. *La Clinica Terapeutica* 2013;**164**(3):e193–8. PUBMED: 23868637]

Calo 2014 *{published data only}*

Calo PG, Pisano G, Medas F, Pittau MR, Gordini L, Demontis R, et al. Identification alone versus intraoperative neuromonitoring of the recurrent laryngeal nerve during thyroid surgery: experience of 2034 consecutive patients. *Journal of Otolaryngology - Head & Neck Surgery* 2014;**43**:16. PUBMED: 24942225]

Calo 2014a *{published data only}*

Calo PG, Medas F, Erdas E, Pittau MR, Demontis R, Pisano G, et al. Role of intraoperative neuromonitoring of recurrent laryngeal nerves in the outcomes of surgery for thyroid cancer. *International Journal of Surgery* 2014;**12 Suppl 1**:S213–7. PUBMED: 24862668]

Cavicchi 2009 *{published data only}*

Cavicchi O, Caliceti U, Fernandez IJ, Macri G, Di Lieto C, Marcantoni A, et al. The value of neurostimulation and

intraoperative nerve monitoring of inferior laryngeal nerve in thyroid surgery. *Otolaryngology - Head and Neck Surgery* 2009;**140**(6):866–70. PUBMED: 19467405]

Chan 2006 ***{published data only}***
Chan WF, Lang BH, Lo CY. The role of intraoperative neuromonitoring of recurrent laryngeal nerve during thyroidectomy: a comparative study on 1000 nerves at risk. *Surgery* 2006;**140**(6):866-72; discussion 872-3. PUBMED: 17188132]

De Falco 2014 ***{published data only}***
De Falco M, Santangelo G, Del Giudice S, Gallucci F, Parmeggiani U. Double probe intraoperative neuromonitoring with a standardized method in thyroid surgery. *International Journal of Surgery* 2014;**12 Suppl 1**: S140–4. PUBMED: 24862671]

Dionigi 2009 ***{published data only}***
Dionigi G, Boni L, Rovera F, Bacuzzi A, Dionigi R. Neuromonitoring and video-assisted thyroidectomy: a prospective, randomized case-control evaluation. *Surgical Endoscopy* 2009;**23**(5):996–1003. PUBMED: 18806939]

Dralle 2004 ***{published data only}***
Dralle H, Sekulla C, Haerting J, Timmermann W, Neumann HJ, Kruse E, et al. Risk factors of paralysis and functional outcome after recurrent laryngeal nerve monitoring in thyroid surgery. *Surgery* 2004;**136**(6): 1310–22. PUBMED: 15657592]

Hei 2016b ***{published data only}***
Hei H, Zhai Y, Qin J, Song Y. Intermittent intraoperative neural monitoring technology in minimally invasive video-assisted thyroidectomy: a preliminary study. *Journal of Investigative Surgery* 2016;**29**(2):93–7. PUBMED: 26891372]

Higgins 2011 ***{published data only}***
Higgins TS, Gupta R, Ketcham AS, Sataloff RT, Wadsworth JT, Sinacori JT. Recurrent laryngeal nerve monitoring versus identification alone on post-thyroidectomy true vocal fold palsy: a meta-analysis. *The Laryngoscope* 2011;**121**(5): 1009–17. PUBMED: 21520117]

Khaled 2012 ***{published data only}***
Khaled AO, Irfan M, Baharudin A, Shahid H. Comparing the morbidity of external laryngeal nerve injury in thyroid surgery with and without identifying the nerve using intraoperative neuromonitoring. *The Medical Journal of Malaysia* 2012;**67**(3):289–92. PUBMED: 23082419]

Lifante 2009 ***{published data only}***
Lifante JC, McGill J, Murry T, Aviv JE, Inabnet WB 3rd. A prospective, randomized trial of nerve monitoring of the external branch of the superior laryngeal nerve during thyroidectomy under local/regional anesthesia and IV sedation. *Surgery* 2009;**146**(6):1167–73. PUBMED: 19958945]

Lombardi 2016 ***{published data only}***
Lombardi CP, Carnassale G, Damiani G, Acampora A, Raffaelli M, De Crea C, et al. "The final countdown": is intraoperative, intermittent neuromonitoring really useful in preventing permanent nerve palsy? Evidence from a meta-analysis. *Surgery* 2016;**160**(6):1693–706. PUBMED: 27566947]

Malik 2016 ***{published data only}***
Malik R, Linos D. Intraoperative neuromonitoring in thyroid surgery: a systematic review. *World Journal of Surgery* 2016;**40**(8):2051–8. PUBMED: 27329143]

Mangano 2014 ***{published data only}***
Mangano A, Lianos GD, Boni L, Kim HY, Roukos DH, Dionigi G. Intraoperative neuromonitoring of the external branch of the superior laryngeal nerve during thyroidectomy: the need for evidence-based data and perioperative technical/technological standardization. *The Scientific World Journal* 2014;**2014**:692365. PUBMED: 25525624]

Masuoka 2015 ***{published data only}***
Masuoka H, Miyauchi A, Higashiyama T, Yabuta T, Fukushima M, Ito Y, et al. Prospective randomized study on injury of the external branch of the superior laryngeal nerve during thyroidectomy comparing intraoperative nerve monitoring and a conventional technique. *Head & Neck* 2015;**37**(10):1456–60. PUBMED: 24844580]

Moris 2014 ***{published data only}***
Moris D, Vernadakis S, Felekouras E. The role of intraoperative nerve monitoring (IONM) in thyroidectomy: where do we stand today?. *Surgical Innovation* 2014;**21**(1): 98–105. PUBMED: 23965591]

Netto 2007 ***{published data only}***
Netto Ide P, Vartanian JG, Ferraz PR, Salgado P, Azevedo JB, Toledo RN, et al. Vocal fold immobility after thyroidectomy with intraoperative recurrent laryngeal nerve monitoring. *Sao Paulo Medical Journal* 2007;**125**(3): 186–90. PUBMED: 17923945]

Parmeggiani 2012 ***{published data only}***
Parmeggiani D, De Falco M, Avenia N, Sanguinetti A, Fiore A, Gubitosi A, et al. NIM vs neurosign in nerve sparing total thyroidectomy. Multicentric experience. *Annali Italiani di Chirurgia* 2012;**83**(3):233–8. PUBMED: 22610121]

Pisanu 2014 ***{published data only}***
Pisanu A, Porceddu G, Podda M, Cois A, Uccheddu A. Systematic review with meta-analysis of studies comparing intraoperative neuromonitoring of recurrent laryngeal nerves versus visualization alone during thyroidectomy. *The Journal of Surgical Research* 2014;**188**(1):152–61. PUBMED: 24433869]

Rulli 2014 ***{published data only}***
Rulli F, Ambrogi V, Dionigi G, Amirhassankhani S, Mineo TC, Ottaviani F, et al. Meta-analysis of recurrent laryngeal nerve injury in thyroid surgery with or without intraoperative nerve monitoring. *Acta Otorhinolaryngologica Italica* 2014;**34**(4):223–9. PUBMED: 25210215]

Stevens 2012 ***{published data only}***
Stevens K, Stojadinovic A, Helou LB, Solomon NP, Howard RS, Shriver CD, et al. The impact of recurrent laryngeal neuromonitoring on multi-dimensional voice outcomes

following thyroid surgery. *Journal of Surgical Oncology* 2012;**105**(1):4–9. PUBMED: 21882195]

Terris 2007 ***{published data only}***

Terris DJ, Anderson SK, Watts TL, Chin E. Laryngeal nerve monitoring and minimally invasive thyroid surgery: complementary technologies. *Archives of Otolaryngology - Head & Neck Surgery* 2007;**133**(12):1254–7. PUBMED: 18086968]

Thomusch 2002 ***{published data only}***

Thomusch O, Sekulla C, Walls G, Machens A, Dralle H. Intraoperative neuromonitoring of surgery for benign goiter. *American Journal of Surgery* 2002;**183**(6):673–8. PUBMED: 12095600]

Uludag 2016 ***{published data only}***

Uludag M, Aygun N, Kartal K, Citgez B, Besler E, Yetkin G, et al. Contribution of intraoperative neural monitoring to preservation of the external branch of the superior laryngeal nerve: a randomized prospective clinical trial. Langenbeck's Archives of Surgery 2016 Dec 29 Epub ahead of print]. DOI: 10.1007/s00423-016-1544-7

Witt 2005 ***{published data only}***

Witt RL. Recurrent laryngeal nerve electrophysiologic monitoring in thyroid surgery: the standard of care?. *Journal of Voice* 2005;**19**(3):497–500. PUBMED: 16102675]

Wojtczak 2016 ***{published data only}***

Wojtczak B, Barczynski M. Intermittent neural monitoring of the recurrent laryngeal nerve in surgery for recurrent goiter. *Gland Surgery* 2016;**5**(5):481–9. PUBMED: 27867862]

Wong 2017 ***{published data only}***

Wong KP, Mak KL, Wong CK, Lang BH. Systematic review and meta-analysis on intra-operative neuro-monitoring in high-risk thyroidectomy. *International Journal of Surgery (London, England)* 2017;**38**:21–30. PUBMED: 28034775]

Zheng 2013 ***{published data only}***

Zheng S, Xu Z, Wei Y, Zeng M, He J. Effect of intraoperative neuromonitoring on recurrent laryngeal nerve palsy rates after thyroid surgery - a meta-analysis. *Journal of the Formosan Medical Association* 2013;**112**(8): 463–72. PUBMED: 24016611]

Additional references

Adam 2016

Adam MA, Thomas S, Youngwirth L, Hyslop T, Reed SD, Scheri RP, et al. Is there a minimum number of thyroidectomies a surgeon should perform to optimize patient outcomes?. *Annals of Surgery* 2017;**265**:402–7. [PUBMED: 26967630]

Al-Qurayshi 2016

Al-Qurayshi Z, Randolph GW, Alshehri M, Kandil E. Analysis of variations in the use of intraoperative nerve monitoring in thyroid surgery. *JAMA Otolaryngology - Head & Neck Surgery* 2016;**142**(6):584–9. [PUBMED: 27101092]

Altman 2003

Altman DG, Bland JM. Interaction revisited: the difference between two estimates. *BMJ* 2003;**326**(7382):219. [PUBMED: 12543843]

Bai 2018

Bai B, Chen W. Protective effects of intraoperative nerve monitoring (IONM) for recurrent laryngeal nerve injury in thyroidectomy: meta-analysis. *Scientific Reports* 2018;**8**(1): 7761. [PUBMED: 29773852]

Barczyński 2009

Barczyński M, Konturek A, Cichoń S. Randomized clinical trial of visualization versus neuromonitoring of recurrent laryngeal nerves during thyroidectomy. *British Journal of Surgery* 2009;**96**(3):240–6. [PUBMED: 19177420]

Bell 2013

Bell ML, McKenzie JE. Designing psycho-oncology randomised trials and cluster randomised trials: variance components and intra-cluster correlation of commonly used psychosocial measures. *Psycho-oncology* 2013;**22**(8): 1738–47.

Borenstein 2017a

Borenstein M, Higgins JP, Hedges LV, Rothstein HR. Basics of meta-analysis: I^2 is not an absolute measure of heterogeneity. *Research Synthesis Methods* 2017;**8**(1):5–18.

Borenstein 2017b

Borenstein M. Prediction intervals. www.meta-analysis.com/prediction (accessed 3 July 2017).

Boutron 2014

Boutron I, Altman DG, Hopewell S, Vera-Badillo F, Tannock I, Ravaud P. Impact of spin in the abstracts of articles reporting results of randomized controlled trials in the field of cancer: the SPIIN randomized controlled trial. *Journal of Clinical Oncology* 2014;**32**(36):4120–6.

Buch 2011

Buch MH, Aletaha D, Emery P, Smolen JS. Reporting of long-term extension studies: lack of consistency calls for consensus. *Annals of the Rheumatic Diseases* 2011;**70**(6): 886–90.

Calò 2014a

Calò PG, Medas F, Erdas E, Pittau MR, Demontis R, Pisano G, et al. Role of intraoperative neuromonitoring of recurrent laryngeal nerves in the outcomes of surgery for thyroid cancer. *International Journal of Surgery (London, England)* 2014;**12 Suppl 1**:S213–7. [PUBMED: 24862668]

Calò 2014b

Calò PG, Pisano G, Medas F, Pittau MR, Gordini L, Demontis R, et al. Identification alone versus intraoperative neuromonitoring of the recurrent laryngeal nerve during thyroid surgery: experience of 2034 consecutive patients. *Journal of Otolaryngology - Head & Neck Surgery* 2014;**43**: 16. [PUBMED: 24942225]

CENTRAL creation details

CENTRAL creation details. www.cochranelibrary.com/central/central-creation (accessed 21 August 2018).

Chandrasekhar 2013
Chandrasekhar SS, Randolph GW, Seidman MD, Rosenfeld RM, Angelos P, Barkmeier-Kraemer J, et al. Clinical practice guideline: improving voice outcomes after thyroid surgery. *Otolaryngology and Head and Neck Surgery* 2013; **148**(6 Suppl):S1–37. [PUBMED: 23733893]

Chen 2015
Chen P, Liang F, Li LY, Zhao GQ. Complications and adverse effects associated with intraoperative nerve monitoring during thyroid surgery under general anesthesia. *Cell Biochemistry and Biophysics* 2015;**71**(2):1029–33.

Chiang 2010
Chiang FY, Lee KW, Chen HC, Chen HY, Lu IC, Kuo WR, et al. Standardization of intraoperative neuromonitoring of recurrent laryngeal nerve in thyroid operation. *World Journal of Surgery* 2010;**34**(2):223–9. [PUBMED: 20020124]

Chiang 2011
Chiang FY, Lu IC, Tsai CJ, Hsiao PJ, Hsu CC, Wu CW. Does extensive dissection of recurrent laryngeal nerve during thyroid operation increase the risk of nerve injury? Evidence from the application of intraoperative neuromonitoring. *American Journal of Otolaryngology* 2011;**32**(6):499–503. [PUBMED: 21306792]

CONSORT 2016
CONSORT Group. The CONSORT statement. www.consort-statement.org (accessed 19 May 2016).

Corbett 2014
Corbett MS, Higgins JP, Woolacott NF. Assessing baseline imbalance in randomised trials: implications for the Cochrane risk of bias tool. *Research Synthesis Methods* 2014; **5**(1):79–85.

Deeks 2017
Deeks JJ, Higgins JPT, Altman DG (editors) on behalf of the Cochrane Statistical Methods Group. Chapter 9: Analysing data and undertaking meta-analyses. In: Higgins JPT, Churchill R, Chandler J, Cumpston MS (editors), Cochrane Handbook for Systematic Reviews of Interventions version 5.2.0 (updated June 2017), Cochrane, 2017. Available from www.training.cochrane.org/handbook.

Deniwar 2015a
Deniwar A, Bhatia P, Kandil E. Electrophysiological neuromonitoring of the laryngeal nerves in thyroid and parathyroid surgery: a review. *World Journal of Experimental Medicine* 2015;**5**(2):120–3. [PUBMED: 25992326]

Deniwar 2015b
Deniwar A, Kandil E, Randolph G. Electrophysiological neural monitoring of the laryngeal nerves in thyroid surgery: review of the current literature. *Gland Surgery* 2015;**4**(5): 368–75. [PUBMED: 26425449]

Dequanter 2015
Dequanter D, Charara F, Shahla M, Lothaire P. Usefulness of neuromonitoring in thyroid surgery. *European Archives of Oto-rhino-laryngology* 2015;**272**(10):3039–43. [PUBMED: 25248913]

Duclos 2011
Duclos A, Lifante JC, Ducarroz S, Soardo P, Colin C, Peix JL. Influence of intraoperative neuromonitoring on surgeons' technique during thyroidectomy. *World Journal of Surgery* 2011;**35**(4):773–8. [PUBMED: 21267565]

Egbewale 2014
Egbewale BE, Lewis M, Sim J. Bias, precision and statistical power of analysis of covariance in the analysis of randomized trials with baseline imbalance: a simulation study. *BMC Medical Research Methodology* 2014;**14:49**:49. DOI: 10.1186/1471-2288-14-49

GRADEproGDT 2015 [Computer program]
McMaster University (developed by Evidence Prime). GRADEproGDT. Version accessed 10 February 2018. Hamilton (ON): McMaster University (developed by Evidence Prime), 2015.

Henry 2017
Henry BM, Graves MJ, Vikse J, Sanna B, Pekala PA, Walocha JA, et al. The current state of intermittent intraoperative neural monitoring for prevention of recurrent laryngeal nerve injury during thyroidectomy: a PRISMA-compliant systematic review of overlapping meta-analyses. *Langenbeck's archives of surgery* 2017;**402**(4):663–73. [PUBMED: 28378238]

Higgins 2002
Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine* 2002;**21**(11):1539–58.

Higgins 2003
Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analysis. *BMJ* 2003;**327** (7414):557–60.

Higgins 2009
Higgins JPT, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 2009; **172**(1):137–59.

Higgins 2011a
Higgins TS, Gupta R, Ketcham AS, Sataloff RT, Wadsworth JT, Sinacori JT. Recurrent laryngeal nerve monitoring versus identification alone on post-thyroidectomy true vocal fold palsy: a meta-analysis. *Laryngoscope* 2011;**121**(5):1009–17. [PUBMED: 21520117]

Higgins 2011b
Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011; Vol. 343:d5928.

Higgins 2011c
Higgins JPT, Deeks JJ, Altman DG (editors). Chapter 16: Special topics in statistics. In: Higgins JPT, Green S (editors), Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org.

Higgins 2017
Higgins JPT, Altman DG, Sterne JAC (editors). Chapter 8: Assessing risk of bias in included studies. In: Higgins JPT, Churchill R, Chandler J, Cumpston MS (editors), Cochrane Handbook for Systematic Reviews of Interventions Version 5.2.0 (updated June 2017), Cochrane, 2017. Available from www.training.cochrane.org/handbook.

Hozo 2005
Hozo SP, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. *BMC Medical Research Methodology* 2005;**5**:13. DOI: 10.1186/1471-2288-5-13

Hróbjartsson 2013
Hróbjartsson A, Thomsen AS, Emanuelsson F, Tendal B, Hilden J, Boutron I, et al. Observer bias in randomized clinical trials with measurement scale outcomes: a systematic review of trials with both blinded and nonblinded assessors. *Canadian Medical Association Journal* 2013;**185** (4):E201–11.

Jones 2015
Jones CW, Keil LG, Holland WC, Caughey MC, Platts-Mills TF. Comparison of registered and published outcomes in randomized controlled trials: a systematic review. *BMC Medicine* 2015;**13**:282. DOI: 10.1186/s12916-015-0520-3

Kirkham 2010
Kirkham JJ, Dwan KM, Altman DG, Gamble C, Dodd S, Smyth R, et al. The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. *BMJ* 2010;**340**:c365. DOI: 10.1136/bmj.c365

Liberati 2009
Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic and meta-analyses of studies that evaluate interventions: explanation and elaboration. *PLoS Medicine* 2009;**6**(7):1–28. DOI: 10.1371/journal.pmed.1000100

Mathieu 2009
Mathieu S, Boutron I, Moher D, Altman DG, Ravaud P. Comparison of registered and published primary outcomes in randomized controlled trials. *JAMA* 2009;**302**(9): 977–84.

Meader 2014
Meader N, King K, Llewellyn A, Norman G, Brown J, Rodgers M, et al. A checklist designed to aid consistency and reproducibility of GRADE assessments: development and pilot validation. *Systematic Reviews* 2014;**3**:82.

Megan 2012
Megan B, Pickering RM, Weatherall M. Design, objectives, execution and reporting of published open-label extension studies. *Journal of Evaluation in Clinical Practice* 2012;**18** (2):209–15.

Pardal-Refoyo 2016
Pardal-Refoyo JL, Ochoa-Sangrador C. Bilateral recurrent laryngeal nerve injury in total thyroidectomy with or without intraoperative neuromonitoring. Systematic review and meta-analysis. *Acta Otorrinolaringologica Espanola* 2016;**67**(2):66–74. [PUBMED: 26025358]

Randolph 2011
Randolph GW, Dralle H, Abdullah H, Barczynski M, Bellantone R, Brauckhoff M, et al. Electrophysiologic recurrent laryngeal nerve monitoring during thyroid and parathyroid surgery: international standards guideline statement. *The Laryngoscope* 2011;**121 Suppl 1**:S1–16. [PUBMED: 21181860]

Review Manager 2014 [Computer program]
Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager 5 (RevMan 5). Version 5.3. Copenhagen: Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

Riley 2011
Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ* 2011;**342**:d549.

Riley 2013
Riley RD, Kauser I, Bland M, Thijs L, Staessen JA, Wang J, et al. Meta-analysis of randomised trials with a continuous outcome according to baseline imbalance and availability of individual participant data. *Statistics in Medicine* 2013;**32** (16):2747–66.

Sanabria 2013
Sanabria A, Ramirez A, Kowalski LP, Silver CE, Shaha AR, Owen RP, et al. Neuromonitoring in thyroidectomy: a meta-analysis of effectiveness from randomized controlled trials. *European archives of oto-rhino-laryngology : official journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS) : affiliated with the German Society for Oto-Rhino-Laryngology - Head and Neck Surgery* 2013;**270**(8):2175–89. [PUBMED: 23681545]

Scherer 2007
Scherer RW, Langenberg P, von EE. Full publication of results initially presented in abstracts. *Cochrane Database of Systematic Reviews* 2007, Issue 2. DOI: 10.1002/14651858.MR000005.pub3

Schünemann 2017
Schünemann HJ, Oxman AD, Higgins JP, Vist GE, Glasziou P, Akl E, et al. on behalf of the Cochrane GRADEing Methods Group and the Cochrane Statistical Methods Group. Chapter 11: Completing 'Summary of findings' tables and grading the confidence in or quality of the evidence. In: Higgins JPT, Churchill R, Chandler J, Cumpston MS (editors), Cochrane Handbook for Systematic Reviews of Interventions version 5.2.0 (updated June2017). Cochrane, 2017. Available from www.training.cochrane.org/handbook.

Sosa 1998
Sosa JA, Bowman HM, Tielsch JM, Powe NR, Gordon TA, Udelsman R. The importance of surgeon experience for clinical and economic outcomes from thyroidectomy. *Annals of Surgery* 1998;**228**(3):320–30. [PUBMED: 9742915]

Sterne 2011
Sterne JA, Sutton AJ, Ioannidis JP, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ* 2011;**343**:d4002.

Sun 2017
Sun W, Liu J, Zhang H, Zhang P, Wang Z, Dong W, et al. A meta-analysis of intraoperative neuromonitoring of recurrent laryngeal nerve palsy during thyroid reoperations. *Clinical Endocrinology* 2017;**87**(5):572–80. [PUBMED: 28585717]

Wong 2006a
Wong SS, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and Embase. *Journal of the Medical Library Association* 2006;**94**(4):451–5.

Wong 2006b
Wong SSL, Wilczynski NL, Haynes RB. Developing optimal search strategies for detecting clinically sound treatment studies in Embase. *Journal of the Medical Library Association* 2006;**94**(1):41–7.

Wood 2008
Wood L, Egger M, Gluud LL, Schulz KF, Jüni P, Altman DG, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ* 2008;**336**(7644):601–5.

Yang 2017
Yang S, Zhou L, Lu Z, Ma B, Ji Q, Wang Y. Systematic review with meta-analysis of intraoperative neuromonitoring during thyroidectomy. *International journal of surgery (London, England)* 2017;**39**:104–13. [PUBMED: 28130189]

Zheng 2015
Zheng H, Jiang L, Wang X, Hu J, Ning J, Wang D, et al. Application experience of intraoperative neuromonitoring in thyroidectomy. *International Journal of Clinical and Experimental Medicine* 2015;**8**(12):22359–64. [PUBMED: 26885214]

References to other published versions of this review

Cirocchi 2016
Cirocchi R, D'Andrea V, Arezzo A, Abraha I, Passera R, Avenia N, et al. Intraoperative neuromonitoring versus visual nerve identification for prevention of recurrent laryngeal nerve injury in adults undergoing thyroid surgery. *Cochrane Database of Systematic Reviews* 2016, Issue 12. DOI: 10.1002/14651858.CD012483

* *Indicates the major publication for the study*

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Barczynski 2009

Methods	**Parallel, randomised controlled clinical trial, randomisation ratio 1:1**
Participants	**Inclusion criteria:** thyroid pathology qualified for first-time bilateral neck surgery. 18-80 years. All genders **Exclusion criteria:** previous thyroid or parathyroid surgery, unilateral thyroid pathology eligible for a minimally invasive approach, mediastinal goitre, preoperative RILN palsy, pregnancy or lactation, age < 18 years, ASA grade 4 and inability to comply with the follow-up protocol **Diagnostic criteria:** not reported
Interventions	**Number of trial centres:** 1 **Treatment before trial:** not reported **Extension period**: no **Description of interventions:** all thyroid operations were performed by 3 experienced endocrine surgeons, who exposed RILNs and performed 2 different interventions for each participant: • **RILN visualisation with associated neuromonitoring**. RILNs were routinely identified by visualisation and had additional nerve monitoring • **RILN visualisation alone**, RILNs were routinely identified by visualisation **IONM technique.** Quote from publication: "Nerve monitoring with the Neurosign® 100 system (Inomed, Teningen, Germany). After identification of the cricoid and thyroid cartilage, the ipsilateral vocal muscle was impaled with the bipolar recording electrode through the cricothyroid ligament. The neutral electrode was placed in the sternocleidomastoid muscle. The proper placement of the electrodes was confirmed by an impedance meter of the circuit in the patient in the final operating position. Before any manipulation of the thyroid gland, the vagus nerve was first dissected over a short stretch to allow for the initial assessment of the indirect stimulation response. A handheld bipolar, concentric stimulating probe was used with a current amplitude of 1 (range 0·5-1·5) mA (depending on the RILN threshold) and 3-Hz impulses of 200 ms each for 1-2 s. The electrical field response of the muscle was documented as an acoustic signal. An attempt was made to identify the RILNs by using the electrode before their visualisation rather than by palpation or surgical dissection. After the removal of the thyroid lobe, both direct stimulation (through an electrode placed on the ipsilateral RILN nerve) and indirect stimulation (through an electrode placed on the ipsilateral vagus nerve) responses were determined. These final stimulation responses were used for predicting the postoperative outcome. The 'laryngeal twitch response' was not evaluated routinely"
Outcomes	Primary outcomes • Permanent RILN palsy: yes • Transient RILN palsy the incidence of the recurrent laryngeal nerve injury (evaluated on 2nd postoperative day and than at 1, 2, 4, 6 and 12 months postoperatively, if paresis was noted on 1st examination) • Health-related quality of life (not evaluated) Secondary outcomes

Barczynski 2009 (*Continued*)

	• The IONM-added value to RILN identification, the value of IONM in prediction of postoperative vocal cords function (intraoperative data compared with observation of vocal cords function postoperatively on the 2nd day postop) • Adverse events other than permanent or transient RILN palsy • Operative time: not evaluated • All-cause mortality: not evaluated • Socioeconomic effects: not evaluated
Study details	**Trial terminated early:** no **Trial ID:** NCT00661024
Publication details	**Language of publication**: English **Funding:** non-commercial funding (Jagiellonian University) **Publication status:** peer-reviewed journal
Stated aim for study	**Quote from publication**: "The aim of this study was to test the hypothesis that identification of the recurrent laryngeal nerve (RLN) during thyroid surgery reduces injury, and that intraoperative nerve monitoring may be of additional benefit"
Notes	-

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	**Quote from publication:** "Randomization was performed by computer-generated permuted block sequencing" **Comment:** unclear block size
Allocation concealment (selection bias)	Low risk	**Quote from publication: "**Randomization was performed by computer-generated permuted block sequencing and allocated using sealed envelopes to be opened in the operating theatre" **Comment:** trial author stated that the envelopes were opaque and sequentially numbered
Blinding of participants and personnel (performance bias) adverse events other than RILN	High risk	**Quote from publication:** "Patients were blinded to their group assignment" **Comment:** trial author stated that the surgical personnel was not blinded
Blinding of participants and personnel (performance bias) all-cause mortality	Low risk	**Comment:** trial author stated that there were no deaths during the trial

Blinding of participants and personnel (performance bias) operative time	High risk	**Quote from publication:** "Patients were blinded to their group assignment" **Comment:** trial author stated that the personnel was not blinded
Blinding of participants and personnel (performance bias) permanent RILN palsy	High risk	**Quote from publication:** "Patients were blinded to their group assignment" **Comment:** trial author stated that the personnel was not blinded
Blinding of participants and personnel (performance bias) transient RILN palsy	High risk	**Quote from publication: "**Patients were blinded to their group assignment" **Comment:** trial author stated that the personnel was not blinded
Blinding of outcome assessment (detection bias) adverse events other than RILN	Unclear risk	**Comment:** trial author stated that there was no blinding for adverse events other than RILN
Blinding of outcome assessment (detection bias) all-cause mortality	Low risk	**Comment:** trial author stated that there were no deaths during the trial
Blinding of outcome assessment (detection bias) operative time	Unclear risk	**Comment:** no blinding of outcome assessment
Blinding of outcome assessment (detection bias) permanent RILN palsy	Low risk	**Comment:** trial author stated that the ENT specialist who performed laryngeal examinations on postoperative day 2 and followed-up participants in case of nerve injury was blinded
Blinding of outcome assessment (detection bias) transient RILN palsy	Low risk	**Comment:** trial author stated that the ENT specialist who performed laryngeal examinations on postoperative day 2 and followed-up participants in case of nerve injury was blinded
Incomplete outcome data (attrition bias) adverse events other than RILN	Low risk	**Comment:** no dropouts (no missing data)
Incomplete outcome data (attrition bias) all-cause mortality	Low risk	**Comment:** trial author stated that there were no deaths during the trial
Incomplete outcome data (attrition bias) operative time	Low risk	**Comment:** no dropouts (no missing data)

Barczynski 2009 (*Continued*)

Incomplete outcome data (attrition bias) permanent RILN palsy	Low risk	**Comment:** no dropouts (no missing data)
Incomplete outcome data (attrition bias) transient RILN palsy	Low risk	**Comment:** no dropouts (no missing data)
Selective reporting (reporting bias)	Low risk	**Comment:** the outcomes reported in the protocol on the ClinicalTrials.gov website and in the published article are the same
Other bias	Unclear risk	**Comment:** no adequate statistical analysis for clustered data (for outcome measures transient and permanent RILN palsy)

Barczynski 2012

Methods	**Parallel, randomised controlled clinical trial, randomisation ratio 1:1**
Participants	**Inclusion criteria:** thyroid pathology qualified for first-time bilateral neck surgery in women with small-moderate sized goitre (< 100 mL in volume). ≥ 18 years. Female gender. **Exclusion criteria:** male gender, previous thyroid or parathyroid surgery, unilateral thyroid pathology eligible for unilateral lobectomy, goitre volume > 100 mL, preoperative RILN palsy, abnormal preoperative voice assessment on GRBAS scale, pregnancy or lactation, age < 18 years, ASA grade 4 and inability to comply with the follow-up protocol **Diagnostic criteria:** not reported
Interventions	**Number of trial centres:** 1 **Treatment before trial:** not reported **Extension period**: no **Description of interventions:** all thyroid operations were performed by 3 experienced endocrine surgeons, who exposed EBSLNs and RILNs and performed 2 different interventions for each participant: • **EBSLN and RILN visualisation associated with neuromonitoring**. The standard practice of attempting to visually identify and preserve the EBSLNs and RILNs continued supported by adjunct of the IONM system • **EBSLN and RILN visualisation alone**. The EBSLNs and RILNs were routinely identified by visualisation alone **IONM technique.** Quote from publication: "The NIM 3.0 system (Medtronic, Jacksonville, FL) was used in all operations of group B patients (105 patients, 210 EBSLNs, and 210 RILNs at risk), with surface electrodes integrated with an endotracheal tube 7.0 in diameter, which was inserted between the vocal folds by an anaesthetist under direct vision during intubation. The standardized technique of neuromonitoring of the RILNs was used, including indirect vagal response evaluation at the beginning and at the end of the operation according to the recommendations formulated recently by the International Intraoperative Monitoring Study Group (Randolph 2011). The nerves were stimulated using a monopolar electrode and the interrupted stimulation technique

Barczynski 2012 (*Continued*)

	at 1 mA, 100 ms impulse duration, and 4 Hz frequency. In case of bifurcated RILN nerves, the assessment included post-stimulation response of each nerve branch based on acoustic evaluation of the signal, EMG response evaluation (latency and amplitude) and the technique of posterior larynx palpation ("laryngeal twitch")"
Outcomes	Primary outcomes • Permanent RILN palsy (evaluated as secondary outcome) • Transient RILN palsy • Health-related quality of life (not evaluated) Secondary outcomes • Adverse events other than permanent or transient RILN palsy: the changes in postoperative voice performance (pre and postoperative assessment, evaluated to 6 months postoperatively): analysis of maximum phonation time, voice level, fundamental frequency and voice quality rating on GRBAS scale • Operative time • All-cause mortality (not evaluated) • Socioeconomic effects (not evaluated)
Study details	**Trial terminated early:** no **Trial ID:** NCT01395134
Publication details	**Language of publication**: English **Funding:** non-commercial funding (Jagiellonian University) **Publication status:** peer-reviewed journal
Stated aim for study	**Quote from publication**: "Intraoperative nerve monitoring (IONM) has gained widespread acceptance as an adjunct to the gold standard of visual nerve identification, and this technique can be used to identify both the recurrent laryngeal nerve (RLN) and the EBSLN. However, it remains unclear whether there is any IONM-added value to the clinical outcome of thyroidectomy in terms of preserved individual voice performance. This study was designed to test that hypothesis."
Notes	-

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	**Quote from publication:** "Randomization was performed by computer and sequencing was based on permuted blocks of 2 and 3 to balance the number of patients in the treatment groups" **Comment:** trial author stated that the block size was adequate
Allocation concealment (selection bias)	Low risk	**Quote from publication:** "Information on the IONM use remained in consecutively numbered and sealed envelopes,

		which were stored in a specific box in the operating room. An envelope containing the allocation was added to the patient's file once he had entered the operating room. In this way, the patient was blinded to the relevant group assignment. Then, the envelope was opened and the surgeon performed the assigned intervention" **Comment:** trial author stated that the envelopes were opaque
Blinding of participants and personnel (performance bias) all-cause mortality	Low risk	**Comment:** trial author stated that there were no deaths during the trial
Blinding of participants and personnel (performance bias) operative time	High risk	**Quote from publication:** "Patients were blinded to the relevant group assignment" **Comment:** trial author stated that the personnel were not blinded
Blinding of participants and personnel (performance bias) permanent RILN palsy	High risk	**Quote from publication:** "Patients were blinded to the relevant group assignment" **Comment:** trial author stated that the personnel were not blinded
Blinding of participants and personnel (performance bias) transient RILN palsy	High risk	**Quote from publication:** "Patients were blinded to the relevant group assignment" **Comment:** trial author stated that the personnel were not blinded
Blinding of outcome assessment (detection bias) all-cause mortality	Low risk	**Comment:** trial author stated that there were no deaths during the trial
Blinding of outcome assessment (detection bias) operative time	Unclear risk	**Comment:** no blinding of outcome assessment
Blinding of outcome assessment (detection bias) permanent RILN palsy	Low risk	**Quote from publication:** "The assessment protocol was strictly followed by the ENT specialist (AH), who was blinded to the patient relevant group assignment"
Blinding of outcome assessment (detection bias) transient RILN palsy	Low risk	**Quote from publication:** "The assessment protocol was strictly followed by the ENT specialist (AH), who was blinded to the patient relevant group assignment"

Barczynski 2012 (*Continued*)

Incomplete outcome data (attrition bias) all-cause mortality	Low risk	**Comment:** trial author stated that there were no deaths during the trial
Incomplete outcome data (attrition bias) operative time	Low risk	**Comment:** no missing data
Incomplete outcome data (attrition bias) permanent RILN palsy	Low risk	**Quote from publication:** "The patients were randomized into two equal-sized groups (n = 105) ... Four group A versus five group B patients were lost to follow-up. Thus, for the final analysis group A consisted of 101 patients ..., whereas group B consisted of 100 patients..." **Comment:** during follow-up a total of 9 participants (4.3%) were lost. The reasons are not provided in the publication but the number of participants lost to follow-up is balanced among treatment arms; it is unlikely that the low attrition rate affected outcome measures
Incomplete outcome data (attrition bias) transient RILN palsy	Low risk	**Quote from publication:** "The patients were randomized into two equal-sized groups (n = 105) ... Four group A versus five group B patients were lost to follow-up. Thus, for the final analysis group A consisted of 101 patients ..., whereas group B consisted of 100 patients..." **Comment:** during follow-up a total of 9 participants were lost (4.3%). The reasons are not provided in the publication but the number of participants lost to follow-up is balanced among treatment arms; it is unlikely that the low attrition rate affected outcome measures
Selective reporting (reporting bias)	Low risk	**Comment:** the outcomes reported in the protocol on the ClinicalTrial.gov website and in the published article are the same
Other bias	Unclear risk	**Comment:** no adequate statistical analysis for clustered data (for outcome measures transient and permanent RILN palsy)

Hei 2016a

Methods	**Parallel randomised controlled clinical trial, randomisation ratio 1:1**
Participants	**Inclusion criteria:** ≥ **1** previous thyroid operation, normal ipsilateral vocal cord function was detected by preoperative laryngoscopy; previous surgical field (either thyroid bed or central neck compartment), as well as ipsilateral RLN, would be exposed during reoperation **Exclusion criteria:** limited movement or paralysis of the ipsilateral vocal cord observed by preoperative laryngoscopy; previous surgical field and ipsilateral RLNs were not exposed during reoperation **Diagnostic criteria:** preoperative neck ultrasound, contrast-enhanced CT scans from the neck to the chest, fibre-optic laryngoscopy, as a gold standard to evaluate the functional integrity of RLNs
Interventions	**Number of trial centres:** 1 **Treatment before trial:** not reported **Extension period**: no **Description of interventions:** all operations were thyroid reoperations performed by 1 experienced thyroid surgeon who had > 20 years' experience with thyroidectomy All the reoperations were performed through the existing skin incision and the RILNs were located and exposed by different methods: • **RILN visualisation associated with neuromonitoring**. The RILNs were located with the assistance of an intermittent IONM system and then fully exposed by the method of nerve mapping • **RILN visualisation alone**. The RILNs were identified visually using different anatomic landmarks, such as Zuckerkandl tubercle, tracheoesophageal groove, and medial aspect of carotid artery. Then the RILNs were dissected upward to the larynx and downward to the thoracic inlet **IONM technique.** Quote from publication: "The RILNs of patients in the NIM group were monitored by the NIM Response 2.0 (Medtronic Xomed, Jacksonville, FL), which had a specific type of endotracheal tube with 4 surface electrodes. After the endotracheal tube was inserted into the trachea by an experienced anaesthetist, a video laryngoscope was used to adjust its position to make sure that the surface electrodes fully contacted with the bilateral vocal cords. This study complied with the standard operating procedures of intermittent IONM. All vagus nerves in the NIM group were dissected out of the carotid sheath for further stimulation. In addition to these, stimulation of the vagus nerve or the RILN was also performed during challenging or concerning maneuvers. Data of peak amplitude and latency were collected and analyzed"
Outcomes	Temporary and permanent RILN paralysis: • temporary RILN paralysis was defined as recovery of RILN function within the first 6 months after thyroid reoperation • permanent RILN paralysis was defined as no recovery of function during this period RILN injuries were divided into 3 groups: • surgeon-related paralysis (defined as unintentional RILN injuries caused by surgical errors, such as ligation, clamping, burn, and so on. RILNs in this group should be neither surrounded by scar tissues nor invaded by tumours) • tumour-related paralysis (defined as intentional injuries because of perineural invasion by tumours or metastatic lymph nodes, and RILNs were partially or completely transected on purpose)

Hei 2016a (*Continued*)

	• scar-related paralysis (defined as RILN injuries because of tissue adhesion)
Study details	**Trial terminated early:** no **Trial ID:** approved by the Ethics Committee of the Affiliated Cancer Hospital of Zhengzhou University and was performed according to the Declaration of Helsinki
Publication details	**Language of publication**: English **Funding:** non-commercial funding (Henan Health Agency; grant no. 112102310259) **Publication status:** peer-reviewed journal
Stated aim for study	**Quote from publication:** "The purpose of this study was to evaluate whether intermittent intraoperative nerve monitoring could reduce the incidence of recurrent laryngeal nerve paralysis in thyroid reoperations"
Notes	-

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	**Quote from publication:** "All enrolled patients signed a written informed consent and then were randomly assigned into either the nerve integrity monitor (NIM) group or the control group" **Comment:** no details
Allocation concealment (selection bias)	Unclear risk	**Quote from publication:** "All enrolled patients signed a written informed consent and then were randomly assigned into either the nerve integrity monitor (NIM) group or the control group" **Comment:** no details
Blinding of participants and personnel (performance bias) permanent RILN palsy	High risk	**Comment:** information about blinding of participants and personnel not provided, personnel probably not blinded
Blinding of participants and personnel (performance bias) transient RILN palsy	High risk	**Comment:** information about blinding of participants and personnel not provided, personnel probably not blinded
Blinding of outcome assessment (detection bias) permanent RILN palsy	Unclear risk	**Comment:** information about blinding of outcome assessment not provided
Blinding of outcome assessment (detection bias) transient RILN palsy	Low risk	**Comment:** information about blinding of outcome assessment not provided

Hei 2016a (*Continued*)

Incomplete outcome data (attrition bias) adverse events other than RILN	Low risk	**Comment:** no dropouts (no missing data)
Incomplete outcome data (attrition bias) permanent RILN palsy	Low risk	**Comment:** no dropouts (no missing data)
Incomplete outcome data (attrition bias) transient RILN palsy	Low risk	**Comment:** no dropouts (no missing data)
Selective reporting (reporting bias)	Unclear risk	**Comment:** protocol number provided, but we could not identify it
Other bias	Unclear risk	**Comment:** no adequate statistical analysis for clustered data (for outcome measures transient and permanent RILN palsy)

Lee 2015

Methods	**Parallel, randomised controlled clinical trial, randomisation ratio not reported**
Participants	**Inclusion criteria:** malignant thyroid nodules **Exclusion criteria:** diffuse toxic goitres (Graves' disease), nodules > 4 cm, recurrent goitres, patients with RILN palsy, and patients who had undergone previous neck surgery **Diagnostic criteria:** the malignant thyroid nodules were confirmed by cytology of fine-needle aspiration
Interventions	**Number of trial centres:** 1 **Treatment before trial:** not reported **Extension period**: no **Description of interventions:** a robotic thyroidectomy, using a standard bilateral axillo-breast approach, was performed with the da Vinci-S Surgical System (Intuitive Surgical, Mountain View, CA) by the same surgeon with > 10 years of experience in robotic surgery. In both groups, the surgeon traced and identified the whole cervical course of the RILN before sealing the inferior thyroid artery: • **RILN visualisation associated with neuromonitoring**. The RILN is identified and exposed solely with the endoscope, giving a magnified view of the nerve. Successively the standard technique for IONM is to stimulate both the vagus nerve and the RILN before thyroid dissection and after complete thyroidectomy • **RILN visualisation alone**. The RILN is identified and exposed solely with the endoscope, giving a magnified view of the nerve **IONM technique.** Quote from publication: "The NIM 2.0 system (Medtronic Xomed Surgical Products, Jacksonville, FL) was used for neuromonitoring of the IONM group according to the international standards guideline. All RLNs in the IONM group can identified by the NIM 2.0 system, so visual identification of RLNs was not an essential surgical process in the IONM group"

Lee 2015 (*Continued*)

Outcomes	• Transient or permanent laryngeal nerve lesions • Voice Handicap Index • Voice Range Profile
Study details	**Trial terminated early:** no **Trial ID:** approved by Korea University Anam Hospital Institutional Review Board (protocol number MD1108)
Publication details	**Language of publication**: English **Funding:** non-commercial funding (Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology; grant no. 012R1A1013413) **Publication status:** peer-reviewed journal
Stated aim for study	**Quote from publication:** "This study evaluates the efficacy of intraoperative neuromonitoring (IONM) on voice performance in robotic thyroidectomy"
Notes	-

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	**Quote from publication:** "The recruited patients were randomized into two groups (IONM and non-IONM) using a block randomization method" **Comment:** not enough details
Allocation concealment (selection bias)	Unclear risk	**Comment:** no details
Blinding of participants and personnel (performance bias) operative time	High risk	**Comment:** information about blinding of participants and personnel not provided, personnel probably not blinded
Blinding of participants and personnel (performance bias) permanent RILN palsy	High risk	**Comment:** information about blinding of participants and personnel not provided, personnel probably not blinded
Blinding of participants and personnel (performance bias) transient RILN palsy	High risk	**Comment:** information about blinding of participants and personnel not provided, personnel probably not blinded
Blinding of outcome assessment (detection bias) adverse events other than RILN	Unclear risk	**Comment:** information about blinding of outcome assessment not provided

Lee 2015 (*Continued*)

Blinding of outcome assessment (detection bias) operative time	Unclear risk	**Comment:** information about blinding of outcome assessment not provided
Blinding of outcome assessment (detection bias) permanent RILN palsy	Unclear risk	**Comment:** information about blinding of outcome assessment not provided
Blinding of outcome assessment (detection bias) transient RILN palsy	Unclear risk	**Comment:** information about blinding of outcome assessment not provided
Incomplete outcome data (attrition bias) adverse events other than RILN	Low risk	
Incomplete outcome data (attrition bias) operative time	Low risk	**Comment:** no dropouts (no missing data)
Incomplete outcome data (attrition bias) permanent RILN palsy	Low risk	**Comment:** no dropouts (no missing data)
Incomplete outcome data (attrition bias) transient RILN palsy	Low risk	**Comment:** no dropouts (no missing data)
Selective reporting (reporting bias)	Unclear risk	**Comment:** protocol number provided, but we could not identify it
Other bias	Low risk	**Comment:** none detected

Sari 2010

Methods	**Parallel, randomised controlled clinical trial, randomisation ratio 1:1**
Participants	**Inclusion criteria:** multinodular goitre, multinodular toxic goitre, Graves' disease, toxic adenoma, solitary adenoma, papillary carcinoma **Exclusion criteria:** the presence of preoperative cord dysfunction, reoperative surgery, retrosternal goitre, monitoring dysfunction (likely electrode displacement) **Diagnostic criteria:** not reported
Interventions	**Number of trial centres:** 1 **Treatment before trial:** not reported **Extension period**: no **Description of interventions:** the same surgeons performed all thyroid operations who exposed RILNs and performed 2 different interventions for each participant: • **RILN visualisation associated with neuromonitoring.** RILNs were routinely identified by visualisation and had additional nerve monitoring • **RILN visualisation alone.** RILNs were routinely identified by visualisation **IONM technique.** Quote from publication: "In IONM group, intubation was per-

	formed without aid of neuromuscular blockade. Endotracheal-based monitoring systems (eg, Medtronic NIM, Jacksonville, FL) are used to monitor the bilateral thyroarytenoid muscles for ongoing real-time EMG activity. Neural stimulation was performed with a disposable nerve with the current set at 1.5 mA. An original EMG signal was obtained from the vagus nerve before identification of RLN. Vagal stimulation is used to assess accuracy of tube placement before dissection near the recurrent laryngeal nerve. The stimulation level was set at 1.5 mA as a starting point and the event threshold at 100 mV. The signal was obtained from the RLN, which was first identified at tracheoesophageal Groove and the RLN was dissected completely from Berry's ligament"
Outcomes	• Identification time of RILN • Operating time • Persistent and transient RILN • Persistent and transient hypoparathyroidism • Identification time of RILN • Mortality • Postoperative complications • Persistent nerve palsy • Persistent hypoparathyroidism • Transient nerve palsy • Transient hypoparathyroidism
Study details	**Trial terminated early:** no **Trial ID:** approved by the ethical committee of the Faculty of Medicine, University of Istanbul
Publication details	**Language of publication**: English **Funding:** no funding reported **Publication status:** peer-reviewed journal
Stated aim for study	**Quote from publication:** "The aim of this clinical trial is to evaluate the effect of the identification time of RILN during thyroidectomy using IONM"
Notes	-

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	**Quote from publication:** "Patients were randomly assigned to have RLNs identified by visualisation alone or with intraoperative nerve monitoring during surgery. Patients were selected according to the number on the random table for two different groups. Randomization was performed by residents"

Allocation concealment (selection bias)	Unclear risk	**Quote from publication:** "Patients were randomly assigned to have RLNs identified by visualisation alone or with intraoperative nerve monitoring during surgery. Patients were selected according to the number on the random table for two different groups. Randomization was performed by residents" **Comment:** no details
Blinding of participants and personnel (performance bias) adverse events other than RILN	High risk	**Comment:** no information about participants and personnel blinding was provided, personnel probably not blinded
Blinding of participants and personnel (performance bias) all-cause mortality	Low risk	**Comment:** no information about participants and personnel blinding was provided, personnel probably not blinded, outcome measure unlikely influenced by lack of blinding
Blinding of participants and personnel (performance bias) operative time	High risk	**Comment:** no information about participants and personnel blinding was provided, personnel probably not blinded
Blinding of participants and personnel (performance bias) permanent RILN palsy	High risk	**Comment:** no information about participants and personnel blinding was provided, personnel probably not blinded
Blinding of participants and personnel (performance bias) transient RILN palsy	High risk	**Comment:** no information about participants and personnel blinding was provided, personnel probably not blinded
Blinding of outcome assessment (detection bias) adverse events other than RILN	Unclear risk	**Comment:** no information was provided about whether the outcome assessor was blinded
Blinding of outcome assessment (detection bias) all-cause mortality	Low risk	**Comment:** no information was provided about whether the outcome assessor was blinded, outcome measure unlikely influence by lack of blinding
Blinding of outcome assessment (detection bias) operative time	Unclear risk	**Comment:** no information was provided about whether the outcome assessor was blinded
Blinding of outcome assessment (detection bias) permanent RILN palsy	Unclear risk	**Comment:** no information was provided whether the outcome assessor was blinded

Sari 2010 (*Continued*)

Blinding of outcome assessment (detection bias) transient RILN palsy	Unclear risk	**Comment:** no information was provided about whether the outcome assessor was blinded
Incomplete outcome data (attrition bias) adverse events other than RILN	Low risk	**Comment:** no dropouts (no missing data)
Incomplete outcome data (attrition bias) all-cause mortality	Low risk	**Comment:** no dropouts (no missing data)
Incomplete outcome data (attrition bias) operative time	Low risk	**Comment:** no dropouts (no missing data)
Incomplete outcome data (attrition bias) permanent RILN palsy	Unclear risk	**Quote from publication:** "Sixteen nerves were excluded from this study (6 nerves had preoperative cord palsy; acoustic signal was not recorded in 10 nerves" **Comment:** 10 nerves of 233 nerves (4.3%) were excluded in the IONM group
Incomplete outcome data (attrition bias) transient RILN palsy	Unclear risk	**Quote from publication:** "Sixteen nerves were excluded from this study (6 nerves had preoperative cord palsy; acoustic signal was not recorded in 10 nerves" **Comment:** 10 nerves of 233 nerves (4.3%) were excluded in the IONM group
Selective reporting (reporting bias)	Unclear risk	**Comment:** no protocol published
Other bias	Unclear risk	**Comment:** no adequate statistical analysis for clustered data (for outcome measures transient and permanent RILN palsy)

Note: where the judgement is 'Unclear' and the description is blank, the trial did not report that particular outcome.
ASA: American Society of Anesthesiology; **CT:** computed tomography; **EBSLN**: external branch of the superior laryngeal nerve; **ENT:** ear, nose and throat; **GRBAS scale**: grade, roughness, breathiness, asthenia, strain scale; **MIVAT**: minimally invasive video-assisted thyroidectomy; **NIM**: nerve integrity monitor; **IONM**: intraoperative nerve monitoring; **RILN**: recurrent inferior laryngeal nerve; **RLN:** recurrent laryngeal nerve; **VAT**: video-assisted thyroidectomy

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Alesina 2012	Not a RCT
Barczynski 2011	Not a RCT
Brauckhoff 2002	Not a RCT
Calo 2013	Not a RCT
Calo 2014	Not a RCT
Calo 2014a	Not a RCT
Cavicchi 2009	The aim of the RCT was to compare neurostimulation with laryngeal palpation versus IONM in the evaluation of RILN safety after thyroidectomy. 1 arm of the trial used IONM, but the other arm did not use visual identification. Neurostimulation with laryngeal palpation is an intermittent monitoring technique for the evaluation of the contraction of the cricoarytenoid muscle (laryngeal twitch) after stimulation of RILN with an electric stimulator probe
Chan 2006	Not a RCT
De Falco 2014	Not a RCT
Dionigi 2009	Quasi-RCT
Dralle 2004	Not a RCT
Hei 2016b	Not a RCT
Higgins 2011	Systematic review and meta-analysis
Khaled 2012	IONM on RILN after thyroidectomy not assessed. Focus was analysis of the efficacy of IONM in preventing an injury of the external branch of the superior laryngeal nerve
Lifante 2009	In this RCT, the impact of using intraoperative neuromonitoring on the recurrent inferior laryngeal nerve after thyroidectomy was not assessed. In effect, the trial focused to analyse the efficacy of intraoperative neuromonitoring in preventing an injury of the external branch of the superior laryngeal nerve
Lombardi 2016	Systematic review and meta-analysis
Malik 2016	Systematic review
Mangano 2014	Systematic review
Masuoka 2015	In this RCT, the impact of using IONM on the recurrent inferior laryngeal nerve after thyroidectomy was not assessed. In effect, the trial focused to analyse the efficacy of intraoperative neuromonitoring in preventing an injury of the external branch of the superior laryngeal nerve

(*Continued*)

Moris 2014	Not a RCT
Netto 2007	Not a RCT
Parmeggiani 2012	RCT used 2 different types of IONM devices in the both arms
Pisanu 2014	Systematic review and meta-analysis
Rulli 2014	Systematic review and meta-analysis
Stevens 2012	Not a RCT
Terris 2007	Not a RCT
Thomusch 2002	Not a RCT
Uludag 2016	IONM on RILN after thyroidectomy not assessed. Focus was analysis of the efficacy of IONM in preventing an injury of the external branch of the superior laryngeal nerve
Witt 2005	Not a RCT
Wojtczak 2016	Narrative review
Wong 2017	Systematic review and meta-analysis
Zheng 2013	Systematic review and meta-analysis

IONM: intraoperative nerve monitoring; **RCT:** randomised controlled trial; **RILN**: recurrent inferior laryngeal nerve

DATA AND ANALYSES

Comparison 1. Intraoperative neuromonitoring versus visual nerve identification only

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Permanent RILN palsy	4	2895	Risk Ratio (M-H, Random, 95% CI)	0.77 [0.33, 1.77]
2 Transient RILN palsy	4	2895	Risk Ratio (M-H, Random, 95% CI)	0.62 [0.35, 1.08]
3 Adverse events other than RILN palsy	2	286	Risk Ratio (M-H, Random, 95% CI)	1.25 [0.45, 3.47]
4 Operative time	4	1488	Mean Difference (IV, Random, 95% CI)	-0.80 [-11.22, 9.62]
5 All-cause mortality	3	1438	Risk Ratio (M-H, Random, 95% CI)	0.0 [0.0, 0.0]

Analysis 1.1. Comparison 1 Intraoperative neuromonitoring versus visual nerve identification only, Outcome 1 Permanent RILN palsy.

Review: Intraoperative neuromonitoring versus visual nerve identification for prevention of recurrent laryngeal nerve injury in adults undergoing thyroid surgery

Comparison: 1 Intraoperative neuromonitoring versus visual nerve identification only

Outcome: 1 Permanent RILN palsy

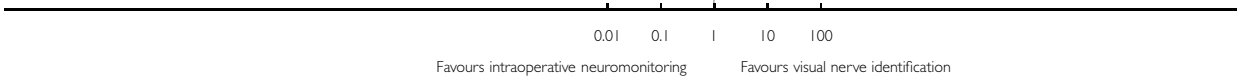
Study or subgroup	Intraoperative neuromonitoring n/N	Visual nerve identification n/N	Risk Ratio M-H,Random,95% CI	Weight	Risk Ratio M-H,Random,95% CI
Barczynski 2009 (1)	8/1000	12/1000		87.6 %	0.67 [0.27, 1.62]
Sari 2010 (2)	0/210	0/199			Not estimable
Barczynski 2012 (3)	0/200	0/202			Not estimable
Hei 2016a (4)	2/41	1/43		12.4 %	2.10 [0.20, 22.26]
Total (95% CI)	**1451**	**1444**		**100.0 %**	**0.77 [0.33, 1.77]**

Total events: 10 (Intraoperative neuromonitoring), 13 (Visual nerve identification)

Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 0.79$, df = 1 (P = 0.37); $I^2 = 0.0\%$

Test for overall effect: Z = 0.62 (P = 0.54)

Test for subgroup differences: Not applicable



(1) Total numbers refer to nerves at risk

(2) Total numbers refer to nerves at risk

(3) Total numbers refer to nerves at risk

(4) Total numbers refer to nerves at risk

Analysis 1.2. Comparison 1 Intraoperative neuromonitoring versus visual nerve identification only, Outcome 2 Transient RILN palsy.

Review: Intraoperative neuromonitoring versus visual nerve identification for prevention of recurrent laryngeal nerve injury in adults undergoing thyroid surgery

Comparison: 1 Intraoperative neuromonitoring versus visual nerve identification only

Outcome: 2 Transient RILN palsy

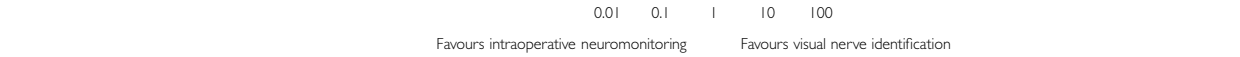
Study or subgroup	Intraoperative neuromonitoring n/N	Visual nerve identification n/N	Risk Ratio M-H,Random,95% CI	Weight	Risk Ratio M-H,Random,95% CI
Barczynski 2009 (1)	19/1000	38/1000		58.3 %	0.50 [0.29, 0.86]
Sari 2010 (2)	3/210	3/199		11.2 %	0.95 [0.19, 4.64]
Barczynski 2012 (3)	3/200	8/202		15.8 %	0.38 [0.10, 1.41]
Hei 2016a (4)	5/41	3/43		14.7 %	1.75 [0.45, 6.85]
Total (95% CI)	**1451**	**1444**		**100.0 %**	**0.62 [0.35, 1.08]**

Total events: 30 (Intraoperative neuromonitoring), 52 (Visual nerve identification)

Heterogeneity: $Tau^2 = 0.06$; $Chi^2 = 3.56$, df = 3 (P = 0.31); $I^2 = 16\%$

Test for overall effect: Z = 1.69 (P = 0.090)

Test for subgroup differences: Not applicable



(1) Total numbers refer to nerves at risk

(2) Total numbers refer to nerves at risk

(3) Total numbers refer to nerves at risk

(4) Total numbers refer to nerves at risk

Analysis 1.3. Comparison 1 Intraoperative neuromonitoring versus visual nerve identification only, Outcome 3 Adverse events other than RILN palsy.

Review: Intraoperative neuromonitoring versus visual nerve identification for prevention of recurrent laryngeal nerve injury in adults undergoing thyroid surgery

Comparison: 1 Intraoperative neuromonitoring versus visual nerve identification only

Outcome: 3 Adverse events other than RILN palsy

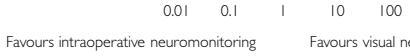
Study or subgroup	Intraoperative neuromonitoring n/N	Visual nerve identification n/N	Risk Ratio M-H,Random,95% CI	Weight	Risk Ratio M-H,Random,95% CI
Sari 2010	11/122	13/114		56.0 %	0.79 [0.37, 1.69]
Lee 2015	9/25	4/25		44.0 %	2.25 [0.80, 6.36]
Total (95% CI)	**147**	**139**		**100.0 %**	**1.25 [0.45, 3.47]**

Total events: 20 (Intraoperative neuromonitoring), 17 (Visual nerve identification)

Heterogeneity: $Tau^2 = 0.33$; $Chi^2 = 2.53$, df = 1 (P = 0.11); $I^2 = 61\%$

Test for overall effect: Z = 0.43 (P = 0.66)

Test for subgroup differences: Not applicable



Analysis 1.4. Comparison 1 Intraoperative neuromonitoring versus visual nerve identification only, Outcome 4 Operative time.

Review: Intraoperative neuromonitoring versus visual nerve identification for prevention of recurrent laryngeal nerve injury in adults undergoing thyroid surgery

Comparison: 1 Intraoperative neuromonitoring versus visual nerve identification only

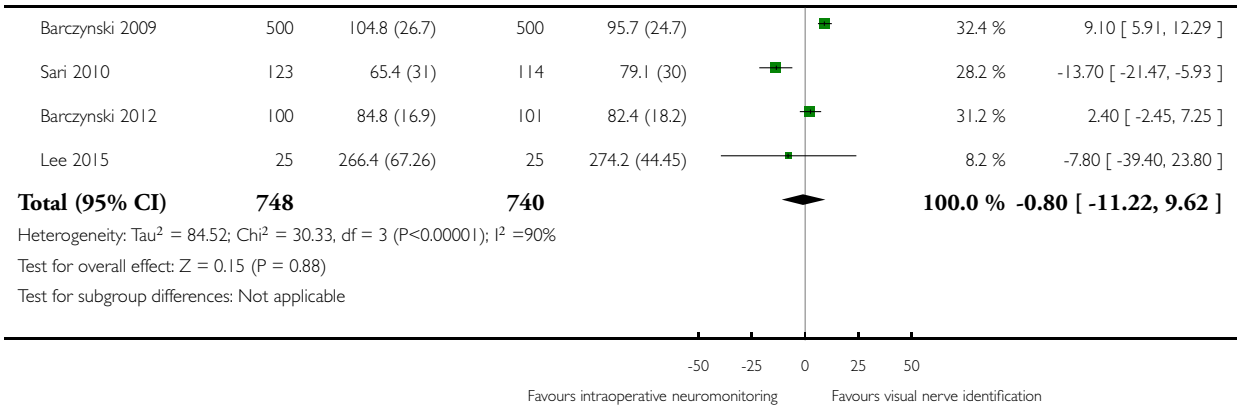
Outcome: 4 Operative time

Study or subgroup	Intraoperative neuromonitoring N	Mean(SD)[min]	Visual nerve identification N	Mean(SD)[min]	Mean Difference IV,Random,95% CI	Weight	Mean Difference IV,Random,95% CI
Barczynski 2009	500	104.8 (26.7)	500	95.7 (24.7)		32.4 %	9.10 [5.91, 12.29]
Sari 2010	123	65.4 (31)	114	79.1 (30)		28.2 %	-13.70 [-21.47, -5.93]
Barczynski 2012	100	84.8 (16.9)	101	82.4 (18.2)		31.2 %	2.40 [-2.45, 7.25]
Lee 2015	25	266.4 (67.26)	25	274.2 (44.45)		8.2 %	-7.80 [-39.40, 23.80]
Total (95% CI)	**748**		**740**			**100.0 %**	**-0.80 [-11.22, 9.62]**

Heterogeneity: $Tau^2 = 84.52$; $Chi^2 = 30.33$, df = 3 ($P<0.00001$); $I^2 = 90\%$

Test for overall effect: Z = 0.15 (P = 0.88)

Test for subgroup differences: Not applicable

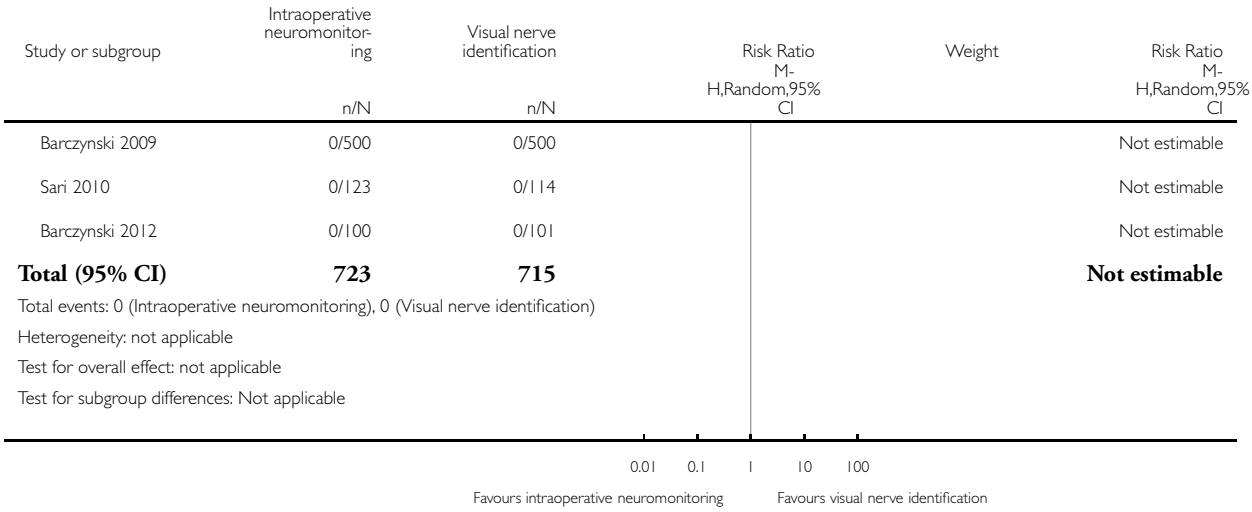


Analysis 1.5. Comparison 1 Intraoperative neuromonitoring versus visual nerve identification only, Outcome 5 All-cause mortality.

Review: Intraoperative neuromonitoring versus visual nerve identification for prevention of recurrent laryngeal nerve injury in adults undergoing thyroid surgery

Comparison: 1 Intraoperative neuromonitoring versus visual nerve identification only

Outcome: 5 All-cause mortality



ADDITIONAL TABLES

Table 1. Overview of trial populations

Trial ID (design)	Intervention(s) and comparator(s)	Description of power and sample size calculation	Screened/ eligible (N)	Randomised (N)	Analysed (N)	Finishing trial (N)	Randomised finishing trial (%)	Follow-up
Hei 2016a (parallel RCT)	I: RILN visualisation with neuromonitoring	"Our study had some limitations. First, this was a small sample study. To draw more persuasive conclusions, at least 434	-	33	33	33	100	6 months

Table 1. Overview of trial populations (*Continued*)

		RLNs in each group are needed to evaluate RLN injury reduction from 10% to 5% with a power of 80% and						
	C: RILN visualisation alone			37	37	37	100	
	total:			**70**	**70**	**70**	**100**	
Lee 2015 (parallel RCT)	I: RILN visualisation with neuromonitoring	"The sample size was estimated based on the principle of detecting a difference of -10 units for VRP and of -5 for VHI between the mean of the IONM and non-IONM groups with a 90% probability at P < .05, using power curve and sample size tools for one-	-	25	25	25	100	12 months
	C: RILN visualisation alone			25	25	25	100	
	total:			**50**	**50**	**50**	**100**	
Barczynski 2012 (parallel RCT)	I: RILN visualisation with neuromonitoring	"The sample size was estimated based on the principle of detecting a 5 % difference in the incidence of primary or secondary out-	517	105	105	100	95.2	6 months

Table 1. Overview of trial populations (*Continued*)

		come measures with a 90 % probability at P < 0.05"						
	C: RILN visualisation alone			105	105	101	96.2	
	total:			**210**	**210**	**201**	**95.7**	
Sari 2010 (parallel RCT)	I: RILN visualisation with neuromonitoring	-	254	123	-[a]	-[a]	-[a]	12 months
	C: RILN visualisation alone			114	-[a]	-[a]	-[a]	
	total:			**236**	-[a]	-[a]	-[a]	
Barczynski 2009 (parallel RCT)	I: RILN visualisation with neuromonitoring	"The sample size was estimated based on the principle of detecting a 2 per cent difference in the incidence of transient RLN injury with a 90 per cent proba-	1488	500	500	500	100	12 months
	C: RILN visualisation alone			500	500	500	100	
	total:			**1000**				
Grand total	**All interventions**			**781**				
	All comparators			777				
	All interventions			**1558**				

Table 1. Overview of trial populations (*Continued*)

	and comparators							

- denotes not reported.

[a] Trial authors did not report the number of participants but the number of nerves

C: comparator; **I**: intervention; **RCT**: randomised controlled trial; **RILN**: recurrent inferior laryngeal nerve

Table 2. Overview of published meta-analyses on intraoperative neuromonitoring versus visual nerve identification only

Author and year of publication	Number of trials included	RCTs included in current review and other published meta-analyses	Quasi-RCTs included in other published meta-analyses	Number of nerves at risk	Permanent RILN palsy	Transient RILN palsy
Bai 2018	34	Sari 2010 Barczynski 2009	Dionigi 2009	58,247	RD −0.0026 (95% CI −0.0039 to −0.0012)	RR 0.71 (95% CI 0.57 to 0.88)
Yang 2017	24	Hei 2016a Sari 2010 Barczynski 2009	Dionigi 2009	17,203	OR 0.78 (95% CI 0.55 to 1.09)	OR 0.76 (95% CI 0.61 to 0.94)
Wong 2017	10[a]	None	None	10,615	OR 1.33 (95% CI 0.94 to 1.88)	OR 1.47 (95% CI 1.07 to 2.00)
Sun 2017	9[b]	None	None	2436	RR 0.426 (95% CI 0.196 to 0.925)	RR 0.607 (95% CI 0.270 to 1.366)
Lombardi 2016	4[c]	Barczynski 2012 Sari 2010 Barczynski 2009	Dionigi 2009	1465[d]	RD 0.00 (95% CI 0.01 to 0.00)	NR
Lombardi 2016	10[e]	None	None	38,820	RR 0.79 (95% CI 0.61to 1.01)	NR
Malik 2016	17	Barczynski 2009	None	44,575	NR	NR
Rulli 2014	8	Barczynski 2009	Dionigi 2009	5257	RR 0.73 (95% CI 0.44 to 1.23)	RR 0.73 (95% CI 0.54 to 0.98)

Table 2. Overview of published meta-analyses on intraoperative neuromonitoring versus visual nerve identification only (*Continued*)

Pisanu 2014	20	Barczynski 2012 Sari 2010 Barczynski 2009	Dionigi 2009	35,513	OR 0.884 (95% CI 0.687 to 1.136)	OR 0.946 (95% CI 0.817 to 1.096)
Sanabria 2013	6	Barczynski 2012 Sari 2010 Barczynski 2009	Dionigi 2009	3064	RD 0 (95% CI -1 to 0)	RD −2 (95% CI −5.1 to 1)
Zheng 2013	14	Sari 2010 Barczynski 2009	None	36,487	OR 0.74 (95% CI 0.59 to 0.92)	OR 0.80 (95% CI 0.65 to 0.99)

[a] High-risk thyroidectomy (reoperation, thyroidectomy for malignancy, thyrotoxicosis or retrosternal goitre)
[b] Thyroid re-operations
[c] Only RCTs
[d] Participants
[e] Only non-RCTs
CI: confidence interval; **EBSLN**: external branch of superior laryngeal; **N**: number; **NAR**: nerves at risk; **NR**: not reported; **OR**: odds ratio; **RCT**: randomized controlled trial; **RD**: risk difference; **RILN**: recurrent inferior laryngeal nerve; **RR:** risk ratio

APPENDICES

Appendix 1. Search strategies

MEDLINE (OvidSP)
1. Recurrent Laryngeal Nerve Injuries/
2. Vocal Cord Paralysis/
3. Recurrent Laryngeal Nerve/
4. Intraoperative Complications/
5. ((vocal or laryngeal) adj3 (nerve? or pals* or paralys* or injur*)).tw
6. rln.tw.
7. or/1-6
8. exp Monitoring, Intraoperative/
9. Electromyography/
10. monitor*.tw.
11. neuromonitor*.tw.
12. (ionm or rlnm).tw.
13. electromyogra*.tw.
14. or/8-13
15. Thyroidectomy/

(*Continued*)

16. Thyroid Diseases/su
17. exp Thyroid Neoplasms/su
18. Thyroid Gland/su
19. ((parathyroid or thyroid) adj3 (surg* or dissect* or resect* or cancer or neoplasm? or operat* or malign*)).tw
20. thyroidectom*.tw.
21. or/15-20
22. 7 and 14 and 21
[23-33: Cochrane Handbook 2008 RCT filter - sensitivity maximizing version]
23. randomized controlled trial.pt.
24. controlled clinical trial.pt.
25. randomi?ed.ab.
26. placebo.ab.
27. drug therapy.fs.
28. randomly.ab.
29. trial.ab.
30. groups.ab.
31. or/23-30
32. exp animals/ not humans/
33. 31 not 32
34. 22 and 33
[35: Wong 2006a*- systematic reviews filter - SensSpec version]*
35. meta analysis.mp,pt. or review.pt. or search*.tw.
36. 22 and 35
37. 34 or 36

Embase (Ovid SP)

1. recurrent laryngeal nerve injury/
2. recurrent laryngeal nerve palsy/
3. recurrent laryngeal nerve/
4. vocal cord paralysis/
5. peroperative complication/
6. ((vocal or laryngeal) adj3 (nerve? or pals* or paralys* or injur*)).tw
7. rln.tw.
8. or/1-7
9. neurophysiological monitoring/
10. neuromonitoring/
11. electromyography/
12. monitor*.tw.
13. neuromonitor*.tw.
14. (ionm or rlnm).tw.
15. electromyogra*.tw.
16. or/9-15
17. exp thyroid surgery/
18. ((parathyroid or thyroid) adj3 (surg* or dissect* or resect* or cancer or neoplasm? or operat* or malign*)).tw
19. thyroidectom*.tw.
20. or/17-19
21. 8 and 16 and 20
[22: Wong 2006b *"sound treatment studies" filter - BS version]*

(*Continued*)

22. random*.tw. or clinical trial*.mp. or exp health care quality/
23. 21 and 22

Cochrane Central Register of Controlled Trials (Cochrane Register of Studies Online)

1. MESH DESCRIPTOR Recurrent Laryngeal Nerve Injuries
2. MESH DESCRIPTOR Vocal Cord Paralysis
3. MESH DESCRIPTOR Recurrent Laryngeal Nerve
4. MESH DESCRIPTOR Intraoperative Complications
5. ((vocal or laryngeal) ADJ3 (nerve? or pals* or paralys* or injur*)):TI,AB,KY
6. rln:TI,AB,KY
7. #1 OR #2 OR #3 OR #4 OR #5 OR #6
8. MESH DESCRIPTOR Monitoring, Intraoperative EXPLODE ALL TREES
9. MESH DESCRIPTOR Electromyography
10. monitor*:TI,AB,KY
11. neuromonitor*:TI,AB,KY
12. (ionm or rlnm):TI,AB,KY
13. electromyogra*:TI,AB,KY
14. #8 OR #9 OR #10 OR #11 OR #12 OR #13
15. MESH DESCRIPTOR Thyroidectomy
16. MESH DESCRIPTOR Thyroid Diseases WITH QUALIFIERS SU
17. MESH DESCRIPTOR Thyroid Neoplasms EXPLODE ALL TREES WITH QUALIFIERS SU
18. MESH DESCRIPTOR Thyroid Gland WITH QUALIFIERS SU
19. ((parathyroid or thyroid) ADJ3 (surg* or dissect* or resect* or cancer or neoplasm? or operat* or malign*)):TI,AB,KY
20. thyroidectom*:TI,AB,KY
21. #15 OR #16 OR #17 OR #18 OR #19 OR #20
22. #7 AND #14 AND #21

WHO ICTRP Search Portal (Standard search)

laryin* AND neuromonitor* OR
vocal AND neuromonitor* OR
rln AND neuromonitor* OR
laryin* AND monitor* OR
vocal AND monitor* OR
rln AND monitor* OR
laryin* AND electromyograph* OR
vocal AND electromyograph* OR
rln AND electromyograph* OR
ionm OR
rlnm

ClinicalTrials.gov (Basic search)

(laryngeal OR vocal OR RLN OR complication OR complications) AND (monitor OR neuromonitor OR monitoring OR neuromonitoring OR electromyography OR electromyographic OR IONM OR RLNM) AND (thyroid OR parathyroid OR thyroidectomy)

Appendix 2. Assessment of risk of bias

'Risk of bias' domains

Random sequence generation (selection bias due to inadequate generation of a randomised sequence)
For each included trial, we described the method used to generate the allocation sequence in sufficient detail to allow an assessment of whether it should produce comparable groups

- Low risk of bias: the trial authors achieved sequence generation using computer-generated random numbers or a random numbers table. Drawing of lots, tossing a coin, shuffling cards or envelopes, and throwing dice are adequate if an independent person performed this who was not otherwise involved in the trial. We considered the use of the minimisation technique as equivalent to being random.
- Unclear risk of bias: insufficient information about the sequence generation process.
- High risk of bias: the sequence generation method was non-random or quasi-random (e.g. sequence generated by odd or even date of birth; sequence generated by some rule based on date (or day) of admission; sequence generated by some rule based on hospital or clinic record number; allocation by judgment of the clinician; allocation by preference of the participant; allocation based on the results of a laboratory test or a series of tests; or allocation by availability of the intervention).

Allocation concealment (selection bias due to inadequate concealment of allocation prior to assignment)
We described for each included trial the method used to conceal allocation to interventions prior to assignment and we assessed whether intervention allocation could have been foreseen in advance of or during recruitment, or changed after assignment

- Low risk of bias: central allocation (including telephone, interactive voice-recorder, web-based and pharmacy-controlled randomisation); sequentially numbered drug containers of identical appearance; sequentially numbered, opaque, sealed envelopes.
- Unclear risk of bias: insufficient information about the allocation concealment.
- High risk of bias: used an open random allocation schedule (e.g. a list of random numbers); assignment envelopes used without appropriate safeguards; alternation or rotation; date of birth; case record number; any other explicitly unconcealed procedure.

We also evaluated trial baseline data to incorporate assessment of baseline imbalance into the 'Risk of bias' judgment for selection bias (Corbett 2014; Egbewale 2014; Riley 2013). Chance imbalances may also affect judgments on the risk of attrition bias. In the case of unadjusted analyses, we distinguished between trials that we rated as being at low risk of bias on the basis of both randomisation methods and baseline similarity, and trials that we judged as being at low risk of bias on the basis of baseline similarity alone (Corbett 2014). We will reclassify judgements of unclear, low or high risk of selection bias as specified in Appendix 5.

Blinding of participants and study personnel (performance bias due to knowledge of the allocated interventions by participants and personnel during the trial)
We evaluated the risk of detection bias separately for each outcome (Hróbjartsson 2013). We noted whether endpoints were self-reported, investigator-assessed or adjudicated outcome measures (see below)

- Low risk of bias: blinding of participants and key study personnel was ensured, and it was unlikely that the blinding could have been broken; no blinding or incomplete blinding, but we judged that the outcome was unlikely to have been influenced by lack of blinding.
- Unclear risk of bias: insufficient information about the blinding of participants and study personnel; the trial does not address this outcome.
- High risk of bias: no blinding or incomplete blinding, and the outcome was likely to have been influenced by lack of blinding; blinding of trial participants and key personnel attempted, but likely that the blinding could have been broken, and the outcome was likely to be influenced by lack of blinding.

Blinding of outcome assessment (detection bias due to knowledge of the allocated interventions by outcome assessment
We evaluated the risk of detection bias separately for each outcome (Hróbjartsson 2013). We noted whether endpoints were self-reported, investigator-assessed or adjudicated outcome measures (see below)

- Low risk of bias: blinding of outcome assessment is ensured, and it was unlikely that the blinding could have been broken; no blinding of outcome assessment, but we judged that the outcome measurement was unlikely to have been influenced by lack of blinding.
- Unclear risk of bias: insufficient information about the blinding of outcome assessors; the trial did not address this outcome.
- High risk of bias: no blinding of outcome assessment, and the outcome measurement was likely to have been influenced by lack of blinding; blinding of outcome assessment, but likely that the blinding could have been broken, and the outcome

measurement was likely to be influenced by lack of blinding.

Incomplete outcome data (attrition bias due to amount, nature or handling of incomplete outcome data)

For each included trial and/or each outcome, we described the completeness of data, including attrition and exclusions from the analyses. We stated whether the trial reported attrition and exclusions, and report the number of participants included in the analysis at each stage (compared with the number of randomised participants per intervention/comparator groups). We also noted if the trial reported the reasons for attrition or exclusion and whether missing data were balanced across groups or were related to outcomes. We considered the implications of missing outcome data per outcome such as high dropout rates (e.g. above 15%) or disparate attrition rates (e.g. difference of 10% or more between trial arms)

- Low risk of bias: no missing outcome data; reasons for missing outcome data unlikely to be related to true outcome (for survival data, censoring unlikely to introduce bias); missing outcome data balanced in numbers across intervention groups, with similar reasons for missing data across groups; for dichotomous outcome data, the proportion of missing outcomes compared with observed event risk was not enough to have a clinically relevant impact on the intervention effect estimate; for continuous outcome data, plausible effect size (mean difference or standardised mean difference) among missing outcomes was not enough to have a clinically relevant impact on observed effect size; appropriate methods, such as multiple imputation, were used to handle missing data.
- Unclear risk of bias: insufficient information to assess whether missing data in combination with the method used to handle missing data were likely to induce bias; the trial did not address this outcome.
- High risk of bias: reason for missing outcome data was likely to be related to true outcome, with either imbalance in numbers or reasons for missing data across intervention groups; for dichotomous outcome data, the proportion of missing outcomes compared with observed event risk enough to induce clinically relevant bias in the intervention effect estimate; for continuous outcome data, plausible effect size (mean difference or standardised mean difference) among missing outcomes enough to induce clinically-relevant bias in observed effect size; 'as-treated' or similar analysis done with substantial departure of the intervention received from that assigned at randomisation; potentially inappropriate application of simple imputation.

Selective reporting (reporting bias due to selective outcome reporting)

We assessed outcome reporting bias by integrating the results of the appendix 'Matrix of trial endpoints (publications and trial documents)' (Boutron 2014; Jones 2015; Mathieu 2009), with those of the appendix 'High risk of outcome reporting bias according to the Outcome Reporting Bias In Trials (ORBIT) classification' (Kirkham 2010). This analysis formed the basis for the judgement of selective reporting

- Low risk of bias: the trial protocol was available and all the trial's prespecified (primary and secondary) outcomes that were of interest to this review were reported in the prespecified way; the study protocol was unavailable, but it was clear that the published reports included all expected outcomes (ORBIT classification).
- Unclear risk of bias: insufficient information about selective reporting.
- High risk of bias: not all the trial's prespecified primary outcomes were reported; one or more primary outcomes were reported using measurements, analysis methods or subsets of the data (e.g. subscales) that were not prespecified; one or more reported primary outcomes were not prespecified (unless clear justification for their reporting was provided, such as an unexpected adverse effect); one or more outcomes of interest in the Cochrane Review were reported incompletely so that we could not enter them in a meta-analysis; the trial report failed to include results for a key outcome that we would expect to have been reported for such a trial (ORBIT classification).

Other bias

- Low risk of bias: the trial appears to be free from other sources of bias.
- Unclear risk of bias: there was insufficient information to assess whether an important risk of bias existed; insufficient rationale or evidence that an identified problem introduced bias.
- High risk of bias: the trial had a potential source of bias related to the specific trial design used; the trial was claimed to be fraudulent; or the trial had some other serious problem.

Appendix 3. Selection bias decisions

Selection bias decisions for trials reporting unadjusted analyses: comparison of results obtained using method details alone with results using method details and trial baseline information[a]

Reported randomisation and allocation concealment methods	'Risk of bias' judgement using methods reporting	Information gained from study characteristics data	'Risk of bias' judgement using baseline information and methods reporting
Unclear methods	Unclear risk	Baseline imbalances present for important prognostic variable (s)	**High risk**
		Groups appear similar at baseline for all important prognostic variables	**Low risk**
		Limited or no baseline details	Unclear risk
Would generate a truly random sample, with robust allocation concealment	Low risk	Baseline imbalances present for important prognostic variable (s)	**Unclear risk**[c]
		Groups appear similar at baseline for all important prognostic variables	Low risk
		Limited baseline details, showing balance in some important prognostic variables[b]	Low risk
		No baseline details	**Unclear risk**
Sequence is not truly random, or allocation concealment is inadequate	High risk	Baseline imbalances present for important prognostic variable (s)	High risk
		Groups appear similar at baseline for all important prognostic variables	**Low risk**
		Limited baseline details, showing balance in some important prognostic variables[b]	**Unclear risk**
		No baseline details	High risk

[a] Taken from Corbett 2014; judgements highlighted in bold indicate situations in which the addition of baseline assessments would change the judgement about risk of selection bias, compared with using methods reporting alone.
[b] Details for the remaining important prognostic variables not reported.
[c] Imbalance identified, which appears likely to be due to chance

Appendix 4. Description of interventions

Trial ID	Intervention	Comparator
Hei 2016a	RILN visualisation + IONM	RILN visualisation
Lee 2015	RILN visualisation + IONM	RILN visualisation
Barczynski 2012	RILN visualisation + IONM	RILN visualisation
Sari 2010	RILN visualisation + IONM	RILN visualisation
Barczynski 2009	RILN visualisation + IONM	RILN visualisation

IONM: intraoperative nerve monitoring; **RILN**: recurrent inferior laryngeal nerve

Appendix 5. Baseline characteristics (I)

Trial ID	Intervention and comparator	Duration of follow-up	Description of participants	Trial period (year to year)	Country	Setting	Ethnic groups (%)	Duration of thyroid disease
Hei 2016a	I: RILN visualisation + IONM	6 months	Participants with thyroid neoplastic (papillary, follicular and medullary carcinoma) (77%) or nontoxic nodular goitre recurrence (23%) after thyroidectomy	January 2012 to August 2014	China	Inpatient	Asian (100)	-
	C: RILN visualisation						Asian (100)	-
Lee 2015	I: RILN visualisation + IONM	12 months	Participants with papillary thyroid carcinoma (100%)	March 2011 to September 2012	Korea	Inpatient	Asian (100)	-

(*Continued*)

	C: RILN visualisation						Asian (100)	-
Barczynski 2012	I: RILN visualisation + IONM	6 months	Participants with Graves' disease (5%), thyroid carcinoma (12%), toxic (15%) and non-toxic (68%) nodular goitre	September 2009 to June 2010	Poland	Inpatient	White (100)	-
	C: RILN visualisation						White (100)	-
Sari 2010	I: RILN visualisation + IONM	12 months	Participants with Graves' disease (9%), toxic adenoma (6%), solitary adenoma (14%), thyroid carcinoma (17%), toxic (14%) and non-toxic (39%) multinodular goitre	September 2007 to September 2009	Turkey	Inpatient	White (100)	-
	C: RILN visualisation						White (100)	-
Barczynski 2009	I: RILN visualisation + IONM	12 months	Participants with Graves' disease (6%), thyroid carcinoma (12%), thyroiditis (2%), toxic (10%) and nontoxic nodular goitre (70%)	January 2006 to June 2007	Poland	Inpatient	White (100)	-
	C: RILN visualisation						White (100)	-

- denotes not reported
C: comparator; **I**: intervention; **IONM**: intraoperative nerve monitoring; **RILN**: recurrent inferior laryngeal nerve

Appendix 6. Baseline characteristics (II)

Trial ID	Intervention and comparator	Sex (female %)	Age (mean/ range years (SD))	BMI (mean kg/ m² (SD))	Type of thyroidectomy (%)	Experience in thyroid surgery	Comedications/co-interventions (% of participants)	Comorbidities (% of participants)
Hei 2016a	I: RILN visualisation + IONM	70	48.3 (9.1)	-	Thyroid reoperation (100)	"1 experienced thyroid surgeon who had more than 20 years' experience with thyroidectomy"	Extended central neck compartment dissection (94)	-
	C: RILN visualisation	84	46.8 (10.6)	-	Thyroid reoperation (100)		Extended central neck compartment dissection (94)	-
Lee 2015	I: RILN visualisation + IONM	92	44.2 (11.9)	-	Robotic thyroidectomy using the bilateral axillo-breast approach (100)	"The same surgeon with more than 10 years of experience in robotic surgery"	-	-
	C: RILN visualisation	76	41.7 (9.0)	-	Robotic thyroidectomy using the bilateral axillo-breast approach (100)		-	-
Barczynski 2012	I: RILN visualisation + IONM	100	50.3 (15.3)	-	Total thyroidectomy (100)	"Three experienced endocrine surgeons"	Central neck compartment dissection (13)	-
	C: RILN visualisation	100	49.7 (14.1)	-	Total thyroidectomy (100)		Central neck compartment dissection (13)	-

(*Continued*)

Sari 2010	I: RILN visualisation + IONM	85	47.2 (14)	26.9 (3)	Total thyroidectomy (19), lobectomy (81)	"The same surgeons in all patients"	-	-
	C: RILN visualisation	80	48.3 (12)	27.3 (3)	Total thyroidectomy (23), lobectomy (77)		-	.
Barczynski 2009	I: RILN visualisation + IONM	76	51.3 (14.4)	-	Total thyroidectomy (76), Dunhill operation (19), bilateral subtotal thyroidectomy (5)	"Three experienced endocrine surgeons"	Central neck compartment dissection (12), lateral neck dissection (3)	-
	C: RILN visualisation	76	51.9 (14.7)	-	Total thyroidectomy (74), Dunhill operation (20), bilateral subtotal thyroidectomy (6)		Central neck compartment dissection (12), lateral neck dissection (2)	-

- denotes not reported
C: comparator; **Dunhill operation**: unilateral thyroid lobectomy and contralateral subtotal thyroid resection; **I**: intervention; **IONM**: intraoperative nerve monitoring; **RILN**: recurrent inferior laryngeal nerve; **SD**: standard deviation

Appendix 7. Matrix of trial endpoints (publications and trial documents)

Trial ID	Endpoints quoted in trial document(s) (ClinicalTrials.gov, FDA/EMA document, manufacturer's website, published design paper)[a]	Trial results available in trial register	Endpoints quoted in publication(s)[b,c]	Endpoints quoted in abstract of publication(s)[b,c]

(*Continued*)

Hei 2016a	N/T		**Primary outcome measures**: temporary and permanent RILN paralysis	**Primary outcome measures**: temporary and permanent RILN paralysis
			Secondary outcome measure: -	**Secondary outcome measure**: -
			Other outcome measure: -	**Other outcome measure**: -
Lee 2015	N/T		**Primary outcome measures**: transient or permanent laryngeal nerve lesions; the Voice Handicap Index and the Voice Range Profile	**Primary outcome measures**: transient or permanent laryngeal nerve lesions, the Voice Handicap Index and the Voice Range Profile
			Secondary outcome measure: -	**Secondary outcome measure**: -
			Other outcome measure: -	**Other outcome measure**: -
Barczynski 2012	**Source:** NCT01395134 **Primary outcome measure**: identification rate of the EBSLN	Yes (last verified: 1 March 2017)	**Primary outcome measure**: identification rate of the EBSLN	**Primary outcome measure**: identification rate of the EBSLN
	Secondary outcome measures: anatomical variability of the external branch of the superior laryngeal nerve according to Cernea classification, incidence of EBSLN and RILN injuries assessed by videostroboscopy (transient and permanent) , changes in postoperative voice performance (pre- and postoperative assessment): analysis of maximum phonation time, voice level, fundamental frequency and voice quality rating on GRBAS scale		**Secondary outcome measure**: incidence of EBSLN and RILN injuries assessed by videostroboscopy (transient, permanent and overall); changes in postoperative voice performance (pre- and postoperative assessment): analysis of maximum phonation time, voice level, fundamental frequency and voice quality rating on GRBAS scale	**Secondary outcome measures**: transient RILN injuries; changes in postoperative voice performance: analysis of maximum phonation time, voice level, fundamental frequency and voice quality rating on GRBAS scale
	Other outcome measure: -		**Other outcome measure**: -	**Other outcome measure**: -

(*Continued*)

	History of changes: 2 documented changes; last change 14 July 2011			
Sari 2010	N/T		**Primary outcome measures**: identification time of RILN, operating time, persistent and transient RILN, persistent and transient hypoparathyroidism	**Primary outcome measures**: identification time of RILN, operating time, postoperative complications
			Secondary outcome measure: -	**Secondary outcome measure**: -
			Other outcome measure: -	**Other outcome measure**: -
Barczynski 2009	**Source:** NCT00661024 **Primary outcome measure**: incidence of the recurrent laryngeal nerve injury (evaluated on 2nd postoperative day and at 1, 2, 4, 6 and 12 months postoperatively, if paresis was noted on first examination)	Yes (last verified: 1 March 2017)	**Primary outcome measure**: transient and permanent RILN injuries	**Primary outcome measure**: transient and permanent RILN injuries
	Secondary outcome measures: IONM-added value to RILN identification, the value of IONM in prediction of postoperative vocal cords function (intraoperative data compared with observation of vocal cords function postoperatively on the 2nd postoperative day)		**Secondary outcome measures**: IONM-added value to RILN identification, value of IONM in prediction of postoperative vocal cords function	**Secondary outcome measure**: -
	Other outcome measure: -		**Other outcome measures**: technical problems and intraoperative complications related to IONM	**Other outcome measure**: -
	History of changes: 2 documented changes; last change 17 April 2008			

- denotes not reported

[a] Trial document(s) refers to all available information from published design papers and sources other than regular publications (e.g. FDA/EMA documents, manufacturer's websites, trials registers).

(*Continued*)

[b]Publication(s) refers to trial information published in scientific journals (primary reference, duplicate publications, companion documents or multiple reports of a primary trial).
[c]Primary and secondary outcomes refer to verbatim specifications in publication/records. Other outcome measures refer to all outcomes not specified as primary or secondary outcome measures
EMA: European Medicines Agency; **EBSLN**: external branch of the superior laryngeal nerve; **FDA**: Food and Drug Administration (US); **GRBAS**: grade, roughness, breathiness, asthenia, strain; **IONM**: intraoperative nerve monitoring; **N/T**: no trial document available; **RILN**: recurrent inferior laryngeal nerve

Appendix 8. High risk of outcome reporting bias according to ORBIT classification

Trial ID	Outcome	High risk of bias (category A)[a]	High risk of bias (category D)[b]	High risk of bias (category E)[c]	High risk of bias (category G)[d]
Hei 2016a	N/A				
Lee 2015	N/A				
Barczynski 2012	N/A				
Sari 2010	N/A				
Barczynski 2009	N/A				

[a]Clear that outcome was measured and analysed; trial report states that outcome was analysed but reports only that result was not significant (Classification 'A', table 2, Kirkham 2010).
[b]Clear that outcome was measured and analysed; trial report states that outcome was analysed but report no results (Classification 'D', table 2, Kirkham 2010).
[c]Clear that outcome was measured but was not necessarily analysed; judgement says likely to have been analysed but not reported because of non-significant results (Classification 'E', table 2, Kirkham 2010).
[d]Unclear whether outcome was measured; not mentioned, but clinical judgement says likely to have been measured and analysed but not reported on the basis of non-significant results (Classification 'G', table 2, Kirkham 2010).
N/A: not applicable; **ORBIT**: Outcome Reporting Bias In Trials

Appendix 9. Definition of endpoint measurement (I)[a]

Trial ID	All-cause mortality	Operative time	Transient RILN palsy	Health-related quality of life	Permanent RILN palsy	Socioeconomic effects
Hei 2016a	N/R	N/R	"If RILN paralysis occurred, laryngoscopy was car-	N/R	Dysfunction was defined as no recovery of function during the	N/R

(*Continued*)

			ried out routinely at 1, 3, and 6 months after operation and at the time that the patients felt that their voice obviously improved“ (IO)		first 6 months after thyroid reoperation (IO)	
Lee 2015	N/R	N/D	”VHI, VRP, and laryngoscopy were used to test voice function before surgery and at 2 weeks, 3 months, and 6 months after the operation“ (IO)	N/R	RILN palsy was considered permanent if it persisted for 12 months (IO)	N/R
Barczynski 2012	N/D	The time from skin incision to skin closure”(IO)	“VSL was performed on day 1 postoperatively; in case of abnormal findings, reevaluation was done at 3 and 6 months postoperatively” (IO)	N/R	Vocal cord paresis for 6 months or more following the operation was regarded as permanent palsy (IO)	N/R
Sari 2010	N/D	“The time from skin preparation to closure of the skin incisions” (IO)	“In cases of dysphonia with vocal cord injury, indirect laryngoscopy was also performed 1 and 6 months later” (IO)	N/R	“Persistent nerve palsy was defined as persistent dysfunction and clinical dysphonia that lasted for 12 months postoperatively” (IO)	N/R
Barczynski 2009	N/D	“The time from skin incision to skin closure” (IO)	“Indirect laryngoscopy by a throat specialist was mandatory before surgery and on day 2 after surgery. In pa-	N/R	“Vocal cord paresis for more than 12 months after the operation was regarded as permanent palsy” (IO)	N/R

(*Continued*)

			tients with RILN paresis, an additional examination was scheduled at 2 weeks and 1, 2, 4, 6 and 12 months after surgery, or until the vocal cord function recovered" (IO)			

[a] In addition to definition of endpoint measurement, description of who measured the outcome (**AO**: adjudicated outcome measurement; **IO**: investigator-assessed outcome measurement; **SO**: self-reported outcome measurement)
N/D: not defined; **N/R**: not reported; **RILN**: recurrent inferior laryngeal nerve; **VHI**: Voice Handicap Index; **VRP**: Voice Range Profile; **VSL**: videostrobolaryngoscopy

Appendix 10. Definition of endpoint measurement (II)[a]

Trial ID	Adverse events other than permanent or transient RILN palsy	Severe/serious adverse events
Hei 2016a	Transient hypoparathyroidism: N/R Permanent hypoparathyroidism: N/R	N/R
Lee 2015	Transient hypoparathyroidism: N/D (IO) Permanent hypoparathyroidism: N/D (IO)	N/R
Barczynski 2012	Transient hypoparathyroidism: N/R Permanent hypoparathyroidism: N/R	N/R
Sari 2010	Transient hypoparathyroidism: N/D (IO) Permanent hypoparathyroidism: N/D (IO)	N/R
Barczynski 2009	Transient hypoparathyroidism: N/R Permanent hypoparathyroidism: N/R	N/R

[a] In addition to definition of endpoint measurement, description who measured the outcome (**AO**: adjudicated outcome measurement; **IO**: investigator-assessed outcome measurement; **SO**: self-reported outcome measurement)
N/D: not defined; **N/R**: not reported; **RILN**: recurrent inferior laryngeal nerve

Appendix 11. Adverse events (I)

Trial ID	Intervention(s) and comparator (s)	Participants included in analysis (N)	Deaths (N)	Deaths (% of participants)	Participants with at least one adverse event (N)	Participants with at least one adverse event (%)	Participants with at least one severe/serious adverse event (N)	Participants with at least one severe/serious adverse event (%)
Hei 2016a	I: RILN visualisation + IONM	33	-	-	-	-	-	-
	C: RILN visualisation	37	-	-	-	-	-	-
Lee 2015	I: RILN visualisation + IONM	25	-	-	9	36	0	0
	C: RILN visualisation	25	-	-	4	12	0	0
Barczynski 2012	I: RILN visualisation + IONM	100	0	0	-	-	-	-
	C: RILN visualisation	101	0	0	-	-	-	-
Sari 2010	I: RILN visualisation + IONM	122	0	0	11	11.1	-	-
	C: RILN visualisation	114	0	0	13	8.8	-	-
Barczynski 2009	I: RILN visualisation + IONM	500	0	0	-	-	-	-
	C: RILN visualisation	500	0	0	-	-	-	-

- denotes not reported
C: comparator; **I**: intervention; **IONM**: intraoperative nerve monitoring; **N**: number of participants; **RILN**: recurrent inferior laryngeal nerve

Appendix 12. Adverse events (II)

Trial ID	Intervention(s) and comparator(s)	Participants included in analysis (N)	Participants discontinuing trial due to an adverse event (N)	Participants discontinuing trial due to an adverse event (%)
Hei 2016a	I: RILN visualisation + IONM	33	0	0
	C: RILN visualisation	37	0	0
Lee 2015	I: RILN visualisation + IONM	25	0	0
	C: RILN visualisation	25	0	0
Barczynski 2012	I: RILN visualisation + IONM	100	0	0
	C: RILN visualisation	101	0	0
Sari 2010	I: RILN visualisation + IONM	122	0	0
	C: RILN visualisation	114	0	0
Barczynski 2009	I: RILN visualisation + IONM	500	0	0
	C: RILN visualisation	500	0	0

- denotes not reported
C: comparator; **I**: intervention; **IONM**: intraoperative nerve monitoring; **RILN**: recurrent inferior laryngeal nerve

Appendix 13. Adverse events (III)

Trial ID	Intervention(s) and comparator(s)	Participants included in analysis (N)	Participants with a specific adverse event (description)	Participants with at least one specific adverse events (N)	Participants with at least one specific adverse event (%)
Hei 2016a	I: RILN visualisation + IONM	33	(1) Permanent RILN palsy (2) Transient RILN palsy	(1) 0 (2) 5	(1) 0 (2) 15.2

(*Continued*)

	C: RILN visualisation	37	(1) Permanent RILN palsy (2) Transient RILN palsy	(1) 0 (2) 3	(1) 0 (2) 8.1
Lee 2015	I: RILN visualisation + IONM	25	(1) Permanent RILN palsy (2) Transient RILN palsy (3) Permanent hypoparathyroidism (4) Transient hypoparathyroidism	(1) 0 (2) 0 (3) 0 (4) 9	(1) 0 (2) 0 (3) 0 (4) 36
	C: RILN visualisation	25	(1) Permanent RILN palsy (2) Transient RILN palsy (3) Permanent hypoparathyroidism (4) Transient hypoparathyroidism	(1) 0 (2) 0 (3) 0 (4) 4	(1) 0 (2) 0 (3) 0 (4) 12
Barczynski 2012	I: RILN visualisation + IONM	100	(1) Permanent RILN palsy (2) Transient RILN palsy	(1) 0 (2) 3	(1) 0 (2) 3
	C: RILN visualisation	101	(1) Permanent RILN palsy (2) Transient RILN palsy	(1) 0 (2) 8	(1) 0 (2) 7.9
Sari 2010	I: RILN visualisation + IONM	122	(1) Permanent RILN palsy (2) Transient RILN palsy (3) Permanent hypoparathyroidism (4) Transient hypoparathyroidism	(1) 0 (2) 3 (3) 0 (4) 11	(1) 0 (2) 2.5 (3) 0 (4) 11.1
	C: RILN visualisation	114	(1) Permanent RILN palsy (2) Transient RILN palsy (3) Permanent hypoparathyroidism	(1) 0 (2) 3 (3) 0 (4) 13	(1) 0 (2) 2.6 (3) 0 (4) 8.8

(*Continued*)

			(4) Transient hypoparathyroidism		
Barczynski 2009	I: RILN visualisation + IONM	500	(1) Permanent RILN palsy (2) Transient RILN palsy	(1) 0 (2) 19	(1) 0 (2) 3.8
	C: RILN visualisation	500	(1) Permanent RILN palsy (2) Transient RILN palsy	(1) 0 (2) 38	(1) 0 (2) 7.6
- denotes not reported **C**: comparator; **I**: intervention; **IONM**: intraoperative nerve monitoring; **RILN**: recurrent inferior laryngeal nerve					

Appendix 14. Survey of trial investigators providing information on included trials

Trial ID	Date trial author contacted	Date trial author replied
Hei 2016a	14 June 2017 and 18 January 2018	No answer
Lee 2015	14 June 2017 and 18 January 2018	No answer
Barczynski 2012	Co-author of this review	
Sari 2010	14 June 2017 and 18 January 2018	No answer
Barczynski 2009	Co-author of this review	

Appendix 15. Checklist to aid consistency and reproducibility of GRADE assessments

	(1) Permanent RILN palsy	(2) Transient RILN palsy	(3) Health-related quality of life	(4) Adverse events other than permanent or transient RILN palsy	(5) Operative time	(6) All-cause mortality	(7) Socioeconoic effects

(*Continued*)

Trial limitations (risk of bias)[a]	Was random sequence generation used (i.e. no potential for selection bias)?	Yes	Yes	N/A	Yes	Yes	Yes	N/A
	Was allocation concealment used (i.e. no potential for selection bias)?	Unclear	Unclear		Unclear	Unclear	Unclear	
	Was there blinding of participants and personnel (i.e. no potential for performance bias) or outcome not likely to be influenced by lack of blinding?	No ()	No ()		Unclear	No ()	Yes	
	Was there blinding of outcome assessment (i.e. no potential for detection bias) or was outcome measurement not likely to be influenced by lack of blinding?	Yes	Unclear		Unclear	Unclear	Yes	

(*Continued*)

	Was an objective outcome used?	Yes	Yes		Yes	Yes	Yes
	Were more than 80% of participants enrolled in trials included in the analysis (i.e. no potential reporting bias)?[e]	Yes	Yes		Yes	Yes	Yes
	Were data reported consistently for the outcome of interest (i.e. no potential selective reporting)?	Yes	Yes		Yes	Yes	Yes
	No other biases reported (i.e. no potential of other bias)?	Unclear	Unclear		Unclear	Unclear	Unclear
	Did the trials end up as scheduled (i.e. not stopped early)?	Yes	Yes		Yes	Yes	Yes
Inconsistency[b]	Point estimates did not vary widely?	Yes	Yes		Yes	Yes	N/A
	To what extent did confidence intervals over-	Substantial	Substantial		Substantial	Substantial	N/A

(*Continued*)

	lap (substantial: all confidence intervals overlap at least one of the included studies point estimate; some: confidence intervals overlap but not all overlap at least one point estimate; no: at least one outlier: where the confidence interval of some of the studies do not overlap with those of most included studies)?							
	Was the direction of effect consistent?	No ()	Yes		No ()	No ()	N/A	
	What was the magnitude of statistical heterogeneity (as measured by I^2) - low (I^2 < 40%), moderate (I^2 40%-60%), high I^2 > 60%)?	Low	Low		High ()	High ()	N/A	

(*Continued*)

	Was the test for heterogeneity statistically significant (P < 0.1)?	Not statistically significant	Not statistically significant	Not statistically significant	Statistically significant ()	N/A
Indirectness	Were the populations in included studies applicable to the decision context?	Highly applicable	Highly applicable	Highly applicable	Highly applicable	Highly applicable
	Were the interventions in the included studies applicable to the decision context?	Highly applicable	Highly applicable	Highly applicable	Highly applicable	Highly applicable
	Was the included outcome not a surrogate outcome?	Yes	Yes	Yes	Yes	Yes
	Was the outcome timeframe sufficient?	Sufficient	Sufficient	Sufficient	Sufficient	Sufficient
	Were the conclusions based on direct comparisons?	Yes	Yes	Yes	Yes	Yes
Imprecision[c]	Was the confidence interval for the pooled estimate not consistent with benefit	No ()	No ()	No ()	No ()	N/A

(*Continued*)

	and harm?						
	What is the magnitude of the median sample size (high: 300 participants, intermediate: 100-300 participants, low: < 100 participants)?[e]	Intermediate	Intermediate		Intermediate	Intermediate	Intermediate
	What was the magnitude of the number of included studies (large: > 10 studies, moderate: 5-10 studies, small: < 5 studies)?[e]	Small ()	Small ()		Small ()	Small ()	Small ()
	Was the outcome a common event (e.g. occurs more than 1/100) ?	No ()	Yes		Yes	N/A	N/A
Publication bias[d]	Was a comprehensive search conducted?	Yes	Yes		Yes	Yes	Yes
	Was grey literature searched?	Yes	Yes		Yes	Yes	Yes
	Were no restrictions applied to	Yes	Yes		Yes	Yes	Yes

(*Continued*)

	study selection on the basis of language?						
	There was no industry influence on studies included in the review?	Yes	Yes		Yes	Yes	Yes
	There was no evidence of funnel plot asymmetry?	N/A	N/A		N/A	N/A	N/A
	There was no discrepancy in findings between published and unpublished trials?	N/A	N/A		N/A	N/A	N/A

[a] Questions on risk of bias are answered in relation to the majority of the aggregated evidence in the meta-analysis rather than to individual trials.
[b] Questions on inconsistency are primarily based on visual assessment of forest plots and the statistical quantification of heterogeneity based on I^2 statistic
[c] When judging the width of the confidence interval it is recommended to use a clinical decision threshold to assess whether the imprecision is clinically meaningful.
[d] Questions address comprehensiveness of the search strategy, industry influence, funnel plot asymmetry and discrepancies between published and unpublished trials.
[e] Depends on the context of the systematic review area.
(): key item for potential downgrading the quality of the evidence (GRADE) as shown in the footnotes of the 'Summary of finding' table(s); N/A: not applicable

CONTRIBUTIONS OF AUTHORS

All review authors read and approved the final review draft.

Roberto Cirocchi (RC): protocol draft, acquisition of trial reports, trial selection, data extraction, data analysis, data interpretation, review of drafts and future review updates

Alberto Arezzo (AA): protocol draft, acquisition of trial reports, trial selection, data extraction, data analysis, data interpretation, review of drafts and future review updates

Vito D'Andrea (VD): protocol draft, acquisition of trial reports, trial selection, data extraction, data analysis, data interpretation, review of drafts and future review updates

Iosief Abraha (IA): protocol draft, acquisition of trial reports, trial selection, data extraction, data analysis, data interpretation, review of drafts and future review updates

Georgi Popivanov (GP): trial selection, data extraction, data analysis, data interpretation, review of drafts and future review updates

Nicola Avenia (NA): protocol draft, acquisition of trial reports, trial selection, data extraction, data analysis, data interpretation, review of drafts and future review updates

Chiara Gerardi (CG): trial selection, data extraction, data analysis, data interpretation, review of drafts and future review updates

Brandon Michael Henry (BMH): trial selection, data extraction, data analysis, data interpretation, review of drafts and future review updates

Justus Randolph (JR): protocol draft, acquisition of trial reports, trial selection, data extraction, data analysis, data interpretation, review of drafts and future review updates

Marcin Barczyński (MB): protocol draft, acquisition of trial reports, trial selection, data extraction, data analysis, data interpretation, review of drafts and future review updates.

DECLARATIONS OF INTEREST

RC: none known

AA: none known

VA: none known

IA: none known

GP: none known

NA: none known

RP: none known

CG: Chiara Gerardi and her institution (along with other collegues) are involved in investigator-initiated trials on the use of HIPEC in the surgical treatment of gastric and colorectal cancer. These studies receive an unconditional grants from a commercial entity.

BMH: none known

JR: none known

MB: was the principal investigator of Barczyński 2009; Barczynski 2012.

SOURCES OF SUPPORT

Internal sources

- Roberto Cirocchi, Italy.

University of Perugia, Italy

External sources

- No sources of support supplied

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

We provided a clearer definition of our primary and secondary outcome measures.

NOTES

Portions of the background and methods sections, the appendices, additional tables and figures 1 to 3 of this review are based on a standard template established by Cochrane Metabolic and Endocrine Disorders.

The Co-ordinating Editor of the CMED Group checked all data extractions and data analyses.