SYSTEMATIC REVIEW

A systematic review and meta-analysis on acoustic voice parameters after "uncomplicated" thyroidectomy

Brian HH LANG[1], MS, FRACS

Carlos KH WONG[2], PhD

Estella PM MA[3], PhD

[1]Department of Surgery, The University of Hong Kong, Hong Kong SAR, China

[2]Department of Family Medicine and Primary Care, University of Hong Kong, 3/F Ap Lei Chau Clinic, 161 Main Street, Ap Lei Chau, Hong Kong

[3]Division of Speech and Hearing Sciences, The University of Hong Kong, Hong Kong SAR, China

Address for Correspondence:

Dr Brian HH Lang

Department of Surgery, Queen Mary Hospital, 102 Pokfulam Road, Hong Kong SAR, China

Tel.: (852) 22554232; Fax No.: (852) 28172291; Email: blang@hku.hk

Keywords: Thyroidectomy; voice; recurrent laryngeal nerve; acoustic analysis; dysphonia

Running title: Voice impairment after thyroidectomy appears temporary

FINANCIAL DISCLOSURE: no financial disclosures

CONFLICT OF INTERESTS: none

Text: 2924 words

ABSTRACT

Background: Post-thyroidectomy voice changes are common even without apparent laryngeal nerve injury. Our study evaluated the impact of open cervical thyroidectomy on five acoustic voice parameters in the early (<3 months) and late (≥3 months) postoperative periods.

Methods: A systematic review was performed to identify studies that quantitatively assessed voice quality by acoustic voice analysis before and after thyroidectomy. Parameters included average fundamental frequency (F_0, Hz), Jitter (%), Shimmer (%), noise-to-harmonic ratio (NHR) and maximum phonation time (MPT) (in secs). Meta-analysis was performed using both fixed- and random-effects models.

Results: Altogether 896 patients were analyzed. Relative to baseline, F_0 significantly worsened in the early period (from 194.9 ± 34.9Hz to 188.0 ± 34.0Hz, p=0.001). This was equivalent to a quarter-tone loss (p=0.004). Shimmer (from 3.15 ± 1.59% to 3.19 ± 1.70%, p=0.040) and MPT (from 17.9 secs to 16.7 secs, p=0.046) also worsened in early period while Jitter and NHR remained unchanged in early and late periods. Males suffered greater deterioration in F_0 (from 120.6 ± 18.8Hz to 111.0 ± 18.5Hz, p=0.048) and in NHR (from 0.12 ± 0.02 to 0.16 ± 0.03, p=0.019) than females in the early period. Four of the five acoustic parameters (F_0, Jitter, Shimmer and NHR) significantly worsened after total thyroidectomy (TT) and not after lesser resection.

Conclusions

F_0, Shimmer and MPT significantly worsened in the early and not in the late postoperative period. F_0 impairment was perceptually significant. Males and those undergoing TT suffered greater voice impairment than their counter-parts during the early period.

INTRODUCTION

Thyroidectomy is one of the most common surgical procedures undertaken worldwide [1]. Although it is associated with a low morbidity when done by high-volume (≥100 cases per year) surgeons, debilitating postoperative voice changes occur not uncommonly without apparent recurrent laryngeal nerve (RLN) injury [2,3]. Presumed causes include injury to external branch of superior laryngeal nerve (EBSLN), orotracheal intubation, surgical adhesions, strap muscles denervation and pain or psychological distress. A recent survey of >200 post-thyroidectomy patients found that over a third of them reported persistent vocal changes and felt these changes more concerning than "traditional" complications like hypocalcemia or unsightly cervical scar [4]. Common voice changes including upper pitch loss, difficulty speaking aloud and voice huskiness could be objectively quantified by a computerized acoustic voice program like the multi-dimensional voice program (MDVP) [2,3]. Five objective parameters have generally been used to measure voice quality and they include average fundamental frequency (F_0), Jitter, Shimmer, Noise-to-Harmonic Ratio (NHR) and maximum phonation time (MPT). F_0 is a measurement of pitch while Jitter and Shimmer measure pitch and amplitude variability, respectively [3,5]. NHR measures the amount of noise present while MPT measures voice capacity (i.e. efficiency of voice production) [3,5]. Although numerous studies reported results of acoustic voice analysis in post-thyroidectomy patients, findings were inconsistent. Furthermore, due to difficulty with patient recruitment and compliance, few studies had sufficient power to quantify longer-term (more than 3 months) acoustic changes. Therefore, it remains unclear whether acoustic voice impairments persist. Our study aimed to summarize the findings from the literature and evaluate the impact of open cervical thyroidectomy on acoustic voice quality parameters in the early (< 3 months) and late periods (≥ 3 months) by pooling data from the current literature.

METHODS

This systematic review and meta-analysis was conducted in accordance with the PRISMA statement [6].

Search strategy

Studies evaluating voice quality or voice-related outcomes after open cervical thyroidectomy were retrieved from the Scopus, Medline (PubMed) and Cochrane Library electronic databases on 2nd December 2014. We used the following free text search terms in "All fields"

#1: 'voice'

#2: 'voice quality'

#3: 'thyroidectomy'

#4: #1 OR #2 AND #3

There was no language restriction or methodological filters. The bibliography of two recent reviews on voice after thyroidectomy was searched for additional relevant references [2,7]

Study selection

All titles identified by the search strategy were independently screened by three authors (BHL,CKW,EPM). Search results were compared, and disagreements were resolved by consensus. Abstracts of potentially relevant titles were then reviewed for eligibility and full-length articles were selected for closer examination. Any studies that quantitatively assessed voice quality by acoustic voice analysis under a controlled environment were considered.

Eligible studies had to report data on at least one of the 5 acoustic parameters (F_0, Jitter, Shimmer, NHR and MPT) generated from MDVP before and after an open cervical thyroidectomy. Controlled environment meant that all audio recordings were done in a quiet room with the patient sitting 15-20cm away from a high quality dynamic microphone and then digitized and analyzed using MDVP (Computerized Speech Lab Model, Kay Elemetrics). For MPT, the longest of three attempts was generally taken. Studies that did not specify the exact time of the postoperative testing or did not separate the results of thyroidectomy from other non-thyroid operations (such as parathyroidectomy) or non-cervical approaches were excluded. Case reports, editorials, expert opinions, reviews without original data and studies on pediatric population were excluded. Multiple reports of the same dataset were assessed and only the most representative or updated report of a study was included.

Data extraction

All data were extracted onto a standardized form. The primary data extracted from each article included: type or design of study, type of surgical services (low- or high-volume center), first authorship, country of origin, year of publication, demographics, extent of surgery (total thyroidectomy (TT) or less-than TT (LTT)), thyroid pathology, type and timing of postoperative acoustic voice assessments. A high-volume center was defined as one that performed over 100 cases per year [8]. Since there was no standardization on the timing of acoustic assessments, they were categorized into early and late periods. Early period was defined as within 3 months of thyroidectomy while late was defined as 3 months or more after thyroidectomy. If two or more assessments were performed within the early or late period (e.g. at 4 and 10 weeks), the average of these assessments was taken. TT comprised near-TT and TT while LTT comprised lobectomy and subtotal thyroidectomy. F_0 was expressed in Hz while Jitter and Shimmer were expressed in %. NHR was a ratio and so no unit was given. MPT was expressed in seconds.

Statistical analysis

Paired *t*-test was used for comparison of five acoustic parameters between pre- and post-operative assessments. To make it more clinically relevant, if significant, comparison was repeated with both preoperative and postoperative F_0 data being converted into semitones by logarithmic scale similar to a piano scale. The conversion equation used was $\log_{10}$ (The highest frequency in Hz / The lowest frequency in Hz) / $\log_{10}2 \times 12$. All the individual outcomes were integrated with the meta-analysis software Review Manager Software 5.0 (Cochrane Collaborative, Oxford, England). Standardized mean differences (SMD) were calculated for significant findings. Results were aggregated and analyzed using both fixed- and random-effect models. Publication bias was estimated by Begg's rank correlation test and Egger's regression test [9,10]. These meta-analyses were conducted using IBM SPSS Version 20.0 for Window and Comprehensive Meta-Analysis Version 2.2.064 (Biostat, Inc.)

RESULTS

Of the 950 titles initially identified from the database search (Figure 1), 36 full-length articles were assessed for inclusion. Twenty were excluded and 16 studies [3,5,11-24] were determined to be eligible for inclusion. Appendix 1 lists these 20 articles [25-44] and the reason for their exclusion. No additional study was found from our search of the two bibliographies in previous reviews [2,7].

Patient selection

Consecutive patients undergoing thyroidectomy were recruited in all the studies examined. All studies considered previous laryngeal diseases, vocal cord paralysis and incomplete follow-up as exclusion criteria. Some studies also imposed an age exclusion criterion [17,20-24].

Baseline characteristics and subjective voice assessments

Table 1 shows a summary of the baseline characteristics and type of subjective voice assessments in the 16 eligible studies (3 retrospective and 13 prospective). All 16 studies were conducted in high-volume centers. In the two robotic-thyroidectomy series, only data from the open group were used [21,23]. A total of 896 patients were analyzed. The cohort comprised predominantly females (83.1%) with a mean (± SD) age of 47.1 ± 24.0 years old. The LTT:TT ratio was 1 to 1.7. Three (18.8%) studies exclusively analyzed TT [13,17,20] while two (12.5%) studies only analyzed LTT [12,19]. Twelve (75.0%) studies performed acoustic voice analysis two or more times in the postoperative period [3,5,11,13,15,17-19,21-24]. Eleven (68.8%) studies analyzed benign and malignant cases together [3,5,11,13-18,20,21] while 3 (18.9%)

included malignant cases only [22-24] and one (6.3%) had benign case only [12]. One (6.3%) did not mention underlying thyroid pathology [19].

Subjective voice changes were commonly reported after thyroidectomy and the incidence ranged between 29.7 to 87.0% depending on definition and timing of the assessment [13,14,17]. "Voice fatigue" was the most common complaint and this was followed by "loss of high pitch" and "difficulty of loud voice" [3,5,12,13]. One study found there was a significant relationship between the extent of surgery and voice change [15] while another study did not [16]. Nevertheless, these voice complaints tended to improve over time [15,19]. Changes in voice quality by perceptual rating using the GRBAS (grade of hoarseness, roughness, breathiness, asthenic and strain) scale was reported in several studies [3,14,18,22,23]. Significant increases in GRBAS score were found in the first 2 weeks but less so after 3-6 months [3,14,18]. Postoperative video-laryngostroboscopy (VLS) was done in all of the studies to exclude any postoperative vocal cord abnormalities and RLN injury but only three studies performed routine postoperative cricothyroid electromyography (CT EMG) in their assessment [5,12,20]. Therefore, subtle injury to EBSLN could not be completely excluded in the other 13 studies. One study found significant correlation between changes in objective parameters and voice symptoms [3].

Table 2 shows a summary of the relative changes in F_0, Jitter and Shimmer before and after thyroidectomy. Both F_0 and Shimmer significantly deteriorated in the early period after thyroidectomy. Figure 2 shows forest plots of F_0 in early and late periods. Pooling data from 12 studies, relative to baseline, F_0 in the early period significantly decreased (from 194.9 ± 34.9 Hz to 188.0 ± 34.0 Hz, p=0.001) with a mean difference of 9.68 (95%CI=6.53-12.82) Hz. Even after excluding an obvious outlier [13] from the pooled data, F_0 in the early period remained significantly decreased [from 193.7 ± 35.9 Hz to 188.9 ± 34.8 Hz, p=0.019). When this decrease was translated into number of semitones, it was equivalent approximately to the loss of a quarter-tone (p=0.004). Potential publication bias was not significant, as confirmed by the Begg analysis (Kendall's tau =-0.364, p=0.119) and the Egger regression test (z=2.227, p=0.153). The decrease in F_0 during the late period also almost reached statistical significance (from 192.3 ± 37.8Hz to 188.8 ± 36.7Hz, p=0.063) with a mean difference of 7.36 (95%CI=3.88-10.84) Hz. Similarly, Shimmer (i.e. amplitude perturbation) significantly worsened in early period (from 3.15 ± 1.59% to 3.19 ± 1.70%, p=0.040) but not in late period (from 3.15 ± 1.47% to 3.07 ± 1.76%, p=0.806). The potential publication bias was not significant, as confirmed by Begg analysis (Kendall's tau = 0.044, p=0.858) and the Egger regression test (z= 0.552, p=0.596). On the other hand, relative to baseline, Jitter remained unchanged during both early and late periods (0.95 ± 0.69% to 0.96 ± 0.74%, p=0.582 and 0.85 ± 0.63% to 0.87 ± 0.70%, p=0.492).

Table 3 shows the relative changes in NHR and MPT before and after thyroidectomy. Relative to baseline, NHR remained unchanged during both early (0.14 ± 0.03 to 0.14 ± 0.33, p=0.600) and late (0.12 ± 0.03 to 0.12 ± 0.03, p=0.464) periods. MPT significantly shortened in early period (17.9 secs to 16.7 secs, p=0.046) and not in late period (p=0.757).

Therefore, overall, only 3 of the 5 acoustic parameters (namely, F_0, shimmer and MPT) significantly deteriorated in the early period and none significantly deteriorated in the late period.

Subgroup analyses of males and females

Four studies reported separate acoustic voice data for males and females. Table 4a compares F_0 in early period between male and female subgroups. In the early period, males suffered a significant decrease in F_0 (from 120.6 ± 18.8 Hz to 111.0 ± 18.5 Hz, *p*=0.048) but not among females (from 205.8 ± 30.0 Hz to 203.0 ± 28.2 Hz, *p*=0.704). Figure 3 compares forest plots of F_0 in males and females in early period. Similarly, males also suffered significant worsening in NHR in the early period (from 0.12 ± 0.02 to 0.16 ± 0.02, *p*=0.019) and not among females (from 0.18 ± 0.03 to 0.12 ± 0.03, *p*=0.140). However, similar trends between males and females were not observed for Jitter, Shimmer and MPT.

Subgroup analyses of TT and LTT

Three studies reported separate acoustic voice data for LTT and TT [3,14,22] and three additional studies exclusively analyzed patients who underwent LTT or TT alone [10,16,19]. They were pooled for the subgroup analyses. Table 4b compares changes in F_0 in early period after TT and LTT. Figure 4 shows forest plots of F_0 in early and late periods after TT. After TT, F_0 significantly reduced in the early (from 202.1 ± 31.5Hz to 185.7 ± 32.4Hz, *p*<0.001) and the late periods (from 193.5 ± 34.2Hz to 187.0 ± 33.8Hz, *p*=0.013) while after LTT, F_0 remained unchanged in both early (from 200.7 ± 40.0 vs. 196.3 ± 40.6Hz, *p*=0.291) and late periods (from 195.8 ± 46.1Hz to 193.6 ± 44.0Hz, *p*=0.805). After TT, Jitter significantly worsened in the early period (0.88 ± 0.57% to 1.18 ± 0.77%) (SMD=0.581, 95%CI=0.270-0.892, *p*<0.001) but not after LTT (1.17 ± 0.88% to 1.16 ± 0.84%) (SMD=0.025, 95%CI=-0.209-0.260, *p*=0.832). Shimmer after TT also significantly worsen in the early period (2.17 ± 0.92% vs. 2.58 ± 1.44%, *p*=0.001) but not after LTT (3.71% ± 1.41% vs. 3.54 ± 1.68%, *p*=0.440). NHR after TT subgroup also significantly worsened in the early postoperative period (0.12 ± 0.03 to 0.14 ± 0.04) (SMD=0.573, 95%CI=0.205-0.941, *p*=0.002) but not after LTT (0.13 ± 0.04 to 0.14 ± 0.05) (SMD=0.046, 95%CI=-0.307-0.409, *p*=0.801).

DISCUSSION

Our data showed that of the 5 acoustic parameters pooled, F_0, Shimmer and MPT were the three parameters which significantly deteriorated only in the early period and not in the late period after thyroidectomy. Both Jitter and NHR did not change both in early and late periods. F_0 in the early period significantly decreased (from 194.9 ± 34.9 Hz to 188.0 ± 34.0 Hz, p=0.001) with a mean difference of 9.6 (95%CI=6.53-12.82) Hz. In terms of clinical significance, this decrease was equivalent to the loss of a quarter-tone. Based on our previous experience, this difference would be considered perceptually significant and could potentially affect normal daily activities, although it was not to the extent of dysphonia and was confined in the early period [45]. Like Shimmer and MPT, the decrease in F_0 became non-significant in the late period (from 192.3 ± 37.8Hz to 188.8 ± 36.7Hz, p=0.063) implying the fact that these measurable acoustic changes were largely temporary and would normalize after 3 months. Our observations were concordant to most previous studies that reported voice quality (including F_0) would normalize within a few months after thyroidectomy [5,13,17,21,23] and lend further support to the notion that voice improves over time in the absence of any detectable RLN injury by VLS.

Other interesting findings included that in the early period, males suffered greater decrease in F_0 (from 120.6 ± 18.8 Hz to 111.0 ± 18.5 Hz, p=0.048) and worsening in NHR (from 0.12 ± 0.02 to 0.16 ± 0.03, p=0.019) than females. Therefore, these may imply thyroidectomy might have a greater impact on voice in male than female patients during early postoperative period. One possible explanation might be because males tended to have larger goiters or more advanced malignancy leading to higher risk of postoperative complications [46]. Nevertheless, this finding appeared discordant to one study that found females were significantly more affected than males after thyroidectomy [16]. However, since data on this subject remained rather limited, more confirmatory studies are required to address this issue.

Similar to previous studies [23], our data showed that relative to baseline, those undergoing TT had a greater chance of suffering from significant poorer voice quality than those undergoing LTT in the early period. Four of the five acoustic parameters significantly worsened in the early period after TT and not after LTT. Significant more Jitter (from 0.88% to 1.18%, $p<0.001$) and higher NHR (0.12 to 0.14, p=0002) were observed after TT but not after LTT. Similarly, in the early period, F_0 and Shimmer significantly worsened after TT (202.1Hz to 185.7Hz, $p<0.001$ and 2.71% to 2.58%, $p<0.001$, respectively) and not after LTT (200.7Hz to 196.3Hz, $p=0.291$ and 3.71% to 3.54%, $p=0.440$, respectively). These findings are concordant with the experience of others that TT is generally associated with increased morbidity than lobectomy because of the greater extent of surgical dissection and higher risk of bilateral subtle neural damage [8]. Given the current controversy on the extent of surgery in early or small differentiated thyroid carcinoma, we believe these findings might lend further support for lobectomy over TT if early voice impairments are of great concern.

However, despite these findings, our data should be interpreted cautiously. First, although all studies performed postoperative VLS to diagnose injury to the RLN, only 3 studies performed routine postoperative CT EMG which is the gold standard test for assessing postoperative ESBLN function [5,12,20]. Therefore, given that so few studies had normal postoperative ESBLN function documented, early acoustic voice impairments could have been a result of unrecognized inadvertent injury to one or both of the ESBLNs. Interestingly, when only the data from the three studies with routine EMG were pooled (*data not shown*), [5,12,20] F_0 was not significantly worse than baseline (195.1Hz to 198.0Hz, respectively, p=0.610). Nevertheless, given the paucity of the data, more studies are required to focus on the issue of whether F_0 remained unchanged when the function of both EBSLNs remained intact. Second, although all studies appeared to be using similar MVDP software from the same manufacturer for acoustic analysis, they might have been subtle differences in the program as there were constant updates with the program over time. Nevertheless, relative to the other parameters, F_0 are generally less susceptible to these updates. Third, our results were subjected to selection biases as those with previous laryngeal diseases, vocal cord paralysis or incomplete follow-up were not pooled for analysis. Fourth, it was not possible to account for the possible effects of baseline characteristic differences as raw data were not available. Lastly, since all studies were conducted in high-volume centers, these results might not be reproducible in lower volume centers [8].

Conclusion

Of the 5 acoustic parameters pooled, F_0, Shimmer and MPT were the three parameters which significantly deteriorated only in the early period and not in the late period after thyroidectomy. On the other hand, both Jitter and NHR did not appear to change significantly in both early and late periods. In terms of the extent of early pitch impairment, it was equivalent to loss of half of a semitone or a quarter-tone on a piano scale. Male patients and those undergoing TT appeared to be suffering from greater voice impairment than their counter-parts in the early postoperative period.

ACKNOWLEDGMENT

None

DISCLOSURE STATEMENT

All authors had nothing to disclose. No competing financial interests exist.

AUTHORS CONTRIBUTIONS

BHH Lang / CKH Wong / E Ma were involved in the review of literature, acquisition of data and drafting and completing the manuscript. BHH Lang / CKH Wong / E Ma were also involved in the review of literature and drafting the manuscript. BHH Lang / CKH Wong / E Ma conceived the study, participated in the co-ordination and the acquisition of data and helped to draft the manuscript. All authors read and approved the final manuscript.

REFERENCES

1. The Surgical Outcomes Monitoring & Improvement Program (SOMIP) report volume 5 (July 2012 – June 2013). Coordinated by the Quality and Safety Division of the Hong Kong Hospital Authority. Available : http://www.ha.org.hk/visitor/ha_index.asp
2. Chandrasekhar SS, Randolph GW, Seidman MD, Rosenfeld RM, Angelos P, Barkmeier-Kraemer J, Benninger MS, Blumin JH, Dennis G, Hanks J, Haymart MR, Kloos RT, Seals B, Schreibstein JM, Thomas MA, Waddington C, Warren B, Robertson PJ. 2013 Clinical practice guideline: improving voice outcomes after thyroid surgery Otolaryngol Head Neck Surg. 2013;148(6):S1-37.
3. Stojadinovic A, Shaha AR, Orlikoff RF, Nissan A, Kornak MF, Singh B, Boyle JO, Shah JP, Brennan MF, Kraus DH. Prospective functional voice assessment in patients undergoing thyroid surgery. Ann Surg 2002;236(6):823-32.
4. Grover G, Sadler GP, Mihai R. Morbidity after thyroid surgery: patient perspective. Laryngoscope 2013;123:2319-23
5. Hong KH, Kim YK. Phonatory characteristics of patients undergoing thyroidectomy without laryngeal nerve injury. Otolaryngol Head Neck Surg 1997;117(4):399-404.
6. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009; 21;6(7):e1000097.
7. Hauch A, Al-Qurayshi Z, Randolph G, Kandil E. Total Thyroidectomy is Associated with Increased Risk of Complications for Low- and High-Volume Surgeons. Ann Surg Oncol.2014;21(12):3844-52

8. Lang BH, Wong CK, Tsang RK, Wong KP, Wong BY. Evaluating the cost-effectiveness of laryngeal examination after elective total thyroidectomy. Ann Surg Oncol 2014;21(11):3548-56.
9. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Biometrics 1994;50(4):1088-101.
10. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;13;315(7109):629-34.
11. Debruyne F, Ostyn F, Delaere P, Wellens W. Acoustic analysis of the speaking voice after thyroidectomy. J Voice 1997;11(4):479-82.
12. Aluffi P, Policarpo M, Cherovac C, Olina M, Dosdegani R, Pia F. Post-thyroidectomy superior laryngeal nerve injury. Eur Arch Otorhinolaryngol 2001;258(9):451-4.
13. Sinagra DL, Montesinos MR, Tacchi VA, Moreno JC, Falco JE, Mezzadri NA, Debonis DL, Curutchet HP. Voice changes after thyroidectomy without recurrent laryngeal nerve injury. J Am Coll Surg 2004;199(4):556-60.
14. de Pedro Netto I, Fae A, Vartanian JG, Barros AP, Correia LM, Toledo RN, Testa JR, Nishimoto IN, Kowalski LP, Carrara-de Angelis E. Voice and vocal self-assessment after thyroidectomy. Head Neck 2006;28(12):1106-14.
15. Soylu L, Ozbas S, Uslu HY, Kocak S. The evaluation of the causes of subjective voice disturbances after thyroid surgery. Am J Surg 2007;194(3):317-22
16. Akyildiz S, Ogut F, Akyildiz M, Engin EZ. A multivariate analysis of objective voice changes after thyroidectomy without laryngeal nerve injury. Arch Otolaryngol Head Neck Surg 2008; 134(6):596-602

17. Lombardi CP, Raffaelli M, De Crea C, D'Alatri L, Maccora D, Marchese MR, Paludetti G, Bellantone R. Long-term outcome of functional post-thyroidectomy voice and swallowing symptoms. Surgery 2009;146(6):1174-81

18. Van Lierde K, D'Haeseleer E, Wuyts FL, Baudonck N, Bernaert L, Vermeersch H. Impact of thyroidectomy without laryngeal nerve injury on vocal quality characteristics: an objective multiparameter approach. Laryngoscope 2010;120(2):338-45.

19. Li C, Tao Z, Qu J, Zhou T, Xia F. A voice acoustic analysis of thyroid adenoma patients after a unilateral thyroid lobectomy. J Voice 2012; 26(1):e23-6.

20. Lombardi CP, D'Alatri L, Marchese MR, Maccora D, Monaco ML, De Crea C, Raffaelli M. Prospective electromyographic evaluation of functional postthyroidectomy voice and swallowing symptoms. World J Surg 2012;36(6):1354-60.

21. Tae K, Kim KY, Yun BR, Ji YB, Park CW, Kim DS, Kim TW. Functional voice and swallowing outcomes after robotic thyroidectomy by a gasless unilateral axillo-breast approach: comparison with open thyroidectomy. Surg Endosc 2012;26(7):1871-7.

22. Lee J, Na KY, Kim RM, Oh Y, Lee JH, Lee J, Lee JS, Kim CH, Soh EY, Chung WY. Postoperative functional voice changes after conventional open or robotic thyroidectomy: a prospective trial. Ann Surg Oncol 2012;19(9):2963-70.

23. Ryu J, Ryu YM, Jung YS, Kim SJ, Lee YJ, Lee EK, Kim SK, Kim TS, Kim TH, Lee CY, Park SY, Chung KW. Extent of thyroidectomy affects vocal and throat functions: a prospective observational study of lobectomy versus total thyroidectomy. Surgery 2013;154(3):611-20

24. Jung SP, Kim SH, Bae SY, Lee SK, Kim S, Choi MY, Kim J, Kim M, Kil WH, Choe JH, Kim JH, Nam SJ, Kim JS. A new subfascial approach in open thyroidectomy: efficacy

for postoperative voice, sensory, and swallowing symptoms. A randomized controlled study. Ann Surg Oncol. 2013;20(12):3869-76.

25. McIvor NP, Flint DJ, Gillibrand J, Morton RP. Thyroid surgery and voice-related outcomes. Aust N Z J Surg 2000;70(3):179-83.

26. Pereira JA, Girvent M, Sancho JJ, Parada C, Sitges-Serra A. Prevalence of long-term upper aerodigestive symptoms after uncomplicated bilateral thyroidectomy. Surgery 2003;133(3):318-22.

27. Musholt TJ, Musholt PB, Garm J, Napiontek U, Keilmann A. Changes of the speaking and singing voice after thyroid or parathyroid surgery. Surgery 2006; 140(6):978-88; discussion 988-9.

28. Lombardi CP, Raffaelli M, D'Alatri L, Marchese MR, Rigante M, Paludetti G, Bellantone R. Voice and swallowing changes after thyroidectomy in patients without inferior laryngeal nerve injuries. Surgery 2006;140(6):1026-32; discussion 1032-4.

29. Meek P, Carding PN, Howard DH, Lennard TW. Voice change following thyroid and parathyroid surgery. J Voice 2008;22(6):765-72.

30. Page C, Zaatar R, Biet A, Strunski V. Subjective voice assessment after thyroid surgery: a prospective study of 395 patients. Indian J Med Sci 2007;61(8):448-54.

31. Lombardi CP, Raffaelli M, D'alatri L, De Crea C, Marchese MR, Maccora D, Paludetti G, Bellantone R. Video-assisted thyroidectomy significantly reduces the risk of early postthyroidectomy voice and swallowing symptoms. World J Surg. 2008;32(5):693-700.

32. Henry LR, Solomon NP, Howard R, Gurevich-Uvena J, Horst LB, Coppit G, Orlikoff R, Libutti SK, Shaha AR, Stojadinovic A. The functional impact on voice of sternothyroid muscle division during thyroidectomy. Ann Surg Oncol. 2008;5(7):2027-33.

33. Ortega J, Cassinello N, Dorcaratto D, Leopaldi E. Computerized acoustic voice analysis and subjective scaled evaluation of the voice can avoid the need for laryngoscopy after thyroid surgery. Surgery 2009;145(3):265-71.
34. Timon CI, Hirani SP, Epstein R, Rafferty MA. Investigation of the impact of thyroid surgery on vocal tract steadiness. J Voice. 2010;24(5):610-3.
35. Henry LR, Helou LB, Solomon NP, Howard RS, Gurevich-Uvena J, Coppit G, Stojadinovic A. Functional voice outcomes after thyroidectomy: an assessment of the Dsyphonia Severity Index (DSI) after thyroidectomy. Surgery 2010;147(6):861-70.
36. Santosh M, Rajashekhar B. Perceptual and acoustic analysis of voice in individuals with total thyroidectomy: pre-post-surgery comparison. Indian J Otolaryngol Head Neck Surg. 2011;63(1):32-9
37. Solomon NP, Awan SN, Helou LB, Stojadinovic A. Acoustic analyses of thyroidectomy-related changes in vowel phonation. J Voice. 2012; 26(6):711-20.
38. Solomon NP, Helou LB, Makashay MJ, Stojadinovic A. Aerodynamic evaluation of the postthyroidectomy voice. J Voice 2012;26(4):454-61.
39. Chun BJ, Bae JS, Chae BJ, Hwang YS, Shim MR, Sun DI. 2012 Early postoperative vocal function evaluation after thyroidectomy using thyroidectomy related voice questionnaire. World J Surg. 2012;36(10):2503-8.
40. Scerrino G, Inviati A, Di Giovanni S, Paladino NC, Di Paola V, Lo Re G, Almasio PL, Cupido F, Gulotta G, Bonventre S. Esophageal motility changes after thyroidectomy; possible associations with postoperative voice and swallowing disorders: preliminary results. Otolaryngol Head Neck Surg. 2013;148(6):926-32.

41. Maeda T, Saito M, Otsuki N, Morimoto K, Takahashi M, Iwaki S, Inoue H, Tomoda C, Miyauchi A, Nibu K. Voice quality after surgical treatment for thyroid cancer. Thyroid. 2013;23(7):847-53.

42. Radowsky JS, Helou LB, Howard RS, Solomon NP, Stojadinovic A. Racial disparities in voice outcomes after thyroid and parathyroid surgery. Surgery. 2013;153(1):103-10.

43. Vicente DA, Solomon NP, Avital I, Henry LR, Howard RS, Helou LB, Coppit GL, Shriver CD, Buckenmaier CC, Libutti SK, Shaha AR, Stojadinovic A. Voice outcomes after total thyroidectomy, partial thyroidectomy, or non-neck surgery using a prospective multifactorial assessment. J Am Coll Surg. 2014;219(1):152-63.

44. Minni A, Ruoppolo G, Barbaro M, Di Lorenzo E, Sementilli G, Bononi M. Long-term (12 to 18 months) functional voice assessment to detect voice alterations after thyroidectomy. Eur Rev Med Pharmacol Sci. 2014;18(12):1704-8.

45. Ma EP, Yiu EM. Multiparametric evaluation of dysphonic severity. J Voice. 2006;20(3):380-90.

46. Lang BH, Yih PC, Lo CY. A review of risk factors and timing for postoperative hematoma after thyroidectomy: is outpatient thyroidectomy really safe? World J Surg. 2012;36(10):2497-502.

Table 1. Patient baseline characteristics and types of voice assessment in the 16 included studies

First author / year	Study design	No. of patients #	Sex (M:F)	Mean age ± SD (range)	Extent of surgery		Postoperative Follow-up visits	Underlying thyroid pathology	Subjective / perceptual assessment
					LTT	TT			
Hong (5) / 1997	PS	54	9 : 45	43 (13-67)	46	8	1m,3m,6m	Benign + Malignant	Symptoms Questionnaire
Debruyne (11) / 1997	PS	47	0 : 47	Not available	47	0	D4,2w,3m	Benign + Malignant	None
Aluffi (12) / 2001	RS	42	6 : 36	52 (33-76)	9	33	12-18m	Benign	Symptoms
Stojadinovic (3) / 2002	PS	50	10 : 40	53 (22-77)	21	29	2w,3m	Benign + Malignant	Symptoms, GRBAS
Sinagra (13)/ 2004	PS	46	3 : 43	43 (20–70)	0	46	2m,4m,6m	Benign + Malignant	-
Netto (14) / 2006	PS	88	11 : 77	46 (17-79)	35	53	2w	Benign + Malignant	GRBAS
Soylu (15)/ 2007	PS	48	6 : 42	47 (19-72)	8	40	D2,3m	Benign + Malignant	Symptoms
Akyildiz (16) / 2008	PS	36	9 : 27	50 (17-74)	27	9	1w	Benign + Malignant	-
Lombardi (17) / 2009	PS	110	13 : 97	46.5 ± 13.2	0	110	3m,12m	Benign + Malignant	Symptoms, questionnaire
Van Lierde (18) / 2010	RS	44	12 : 32	49 (17-72)	7	37	1w,6w,3m	Benign + Malignant	GRBAS, questionnaire
Li (19)/ 2012	RS	32	8 : 24	48 ± 9	32	0	1w,1m,3m	Not available	Symptoms
Lombardi (20) / 2012	PS	32	7 : 25	44.5 ± 10.7	0	32	3m	Benign + Malignant	Symptoms, questionnaire
Tae (21) / 2012	PS	61	0 : 61	54.4 ± 10.6	10	51	1w,1m,3m,6m	Benign + Malignant	Symptoms, questionnaire
Lee (22) / 2012	PS	46	6 : 40	38.5 ± 68	11	35	1w,3m	Malignant	GRBAS

Ryu (23) / 2013	PS	74	33 : 41	44.8 ± 8.5* 50.8 ± 11.3*	33	41	1m,6m,12m	Malignant	Symptoms, GRBAS
Jung (24) / 2013	PS	86	12 : 30	48.0 ± 8.7^	22	20	2w,3m	Malignant	Questionnaire
			6 : 38	51.8 ± 8.7^	22	22			
Overall	**-**	**896**	**151 : 745**	**47.1 ± 24.0**	**330**	**566**	**-**	**-**	**-**

Abbreviations: PS = prospective study; RS = retrospective study; LTT = less-than total thyroidectomy; TT = total thyroidectomy; D = Day; w = week; m = month; GRBAS scale = grade, rough, breathy, asthenic, strained

after excluding those with incomplete follow-up or postoperative vocal cord paresis

*data for LTT and TT

^data for subplatysmal and subfascial groups

Table 2. Relative changes in mean fundamental frequency, jitter (or pitch perturbation) and shimmer (or amplitude perturbation) by acoustic voice analysis <3-month (<3m) and ≥3-month (≥3m) after thyroidectomy.

First author / year	Subgroups available	Mean fundamental frequency (F_0) (Hz)			Jitter or pitch perturbation (%)			Shimmer or amplitude perturbation (%)		
		Baseline	<3m	≥3m	Baseline	<3m	≥3m	Baseline	<3m	≥3m
Hong (5) / 1997	No	207 ± 47	196 ± 35	209 ± 45.4	0.44 ± 0.10	0.41 ± 0.06	0.44 ± 0.13	-	-	-
Debruyne (11) / 1997	No	210.4 ± 24.09	205.4 ± 19.7	-	1.14 ± 0.89	1.13 ± 0.89	-	-	-	-
Aluffi (12) / 2001	No	-	-	216 ± 48.0	0.64	-	-	3.25	-	-
Stojadinovic (3) / 2002	LTT (n=21)	169.23 ± 37.90	170.82 ± 44.40	170.35 ± 40.20	0.55 ± 0.23	0.56 ± 0.46	0.62 ± 0.45	1.90 ± 0.93	1.65 ± 0.54	1.95 ± 0.98
	TT (n=29)	165.60 ± 40.76	168.07 ± 45.50	166.27 ± 41.20	0.68 ± 0.39	0.71 ± 0.40	0.59 ± 0.28	1.74 ± 0.68	2.02 ± 0.87	1.86 ± 0.71
Sinagra (13) / 2004	No	211 ± 17	176.00 ± 19.00	192.00 ± 13.58	-	-	-	0.44 ± 0.18	0.82 ± 0.31	0.56 ± 0.26
Netto (14) / 2006	Males (n=11)	114.2 ± 23.0	110.30 ± 7.50	-	0.4 ± 0.2	0.60 ± 0.30	-	3.7 ± 1.8	3.80 ± 1.40	-
	Females (n=77)	196.8 ± 20.6	197.80 ± 27.10	-	0.5 ± 0.3	0.50 ± 0.30	-	2.8 ± 0.8	3.10 ± 1.40	-
Soylu (15) / 2007	LTT (n=8)	236.54 ± 69.38	217.88 ± 71.80	225.54 ± 57.10	0.605 ± 0.39	0.77 ± 0.56	0.59 ± 0.36	2.55 ± 0.54	2.86 ± 0.85	2.51 ± 0.49
	TT (n=40)	237.71 ± 32.03	221.64 ± 26.30	213.95 ± 35.90	0.397 ± 0.15	0.70 ± 0.16	0.66 ± 0.46	2.44 ± 0.54	2.47 ± 0.65	2.41 ± 0.67
Akyildiz (16) / 2008	Males (n=9)	129.26 ± 17.9	118.06 ± 28.20	-	0.70 ± 0.25	0.91 ± 0.81	-	3.71 ± 1.21	5.00 ± 2.33	-
	Females (n=27)	220.43 ± 31.13	215.80 ± 24.70	-	1.45 ± 1.03	0.99 ± 0.76	-	6.96 ± 4.42	4.25 ± 2.69	-
Lombardi (17)	No	186.8 ±	-	185.00 ±	0.5 ± 0.3	-	0.60 ±	3.3 ± 0.8	-	3.90 ±

/ 2009		38.7		36.35			0.38			1.90
Van Lierde (16) / 2010	Males (n=12)	115.3 ± 18.13	104.75 ± 17.62	117.40 ± 8.88	1.0 ± 0.61	1.25 ± 0.79	0.99 ± 0.81	2.5 ± 0.94	2.75 ± 1.84	1.90 ± 0.64
	Females (n=32)	191.7 ± 25.49	189.10 ± 27.30	189.20 ± 28.50	1.0 ± 0.59	1.19 ± 0.67	1.11 ± 0.91	2.5 ± 0.97	2.65 ± 1.28	2.50 ± 1.02
Li (19) / 2012	Males (n=8)	127.65 ± 13.60	117.43 ± 16.74	126.58 ± 19.30	1.22 ± 0.85	1.29 ± 0.99	0.82 ± 0.99	4.60 ± 1.59	5.23 ± 2.04	3.86 ± 2.14
	Females (n=24)	236.81 ± 52.55	231.80 ± 44.75	242.26 ± 37.90	1.65 ± 1.14	1.27 ± 1.21	0.64 ± 0.73	4.92 ± 1.63	4.04 ± 1.89	3.38 ± 2.69
Lombardi (20) / 2012	No	175.0 ± 30.0	-	179.30 ± 32.70	0.63 ± 0.22	-	0.58 ± 0.38	4.63 ± 0.74	-	3.98 ± 1.25
Tae (21) / 2012	No	207.9 ± 37.0	201.40 ± 34.90	202.15 ± 32.89	1.31 ± 0.89	1.10 ± 0.70	1.13 ± 1.56	3.98 ± 2.51	3.80 ± 2.17	3.37 ± 2.61
Lee (22) / 2012	No	187 ± 40.0	189.70 ± 45.00	190.30 ± 45.00	1.37 ± 0.40	1.38 ± 0.45	1.63 ± 0.45	2.90 ± 1.0	3.26 ± 1.20	3.73 ± 1.30
Ryu (23) / 2013	LTT (n=33)	189.6 ± 44.0	187.70 ± 50.50	181.45 ± 50.76	1.39 ± 0.97	1.42 ± 0.59	1.27 ± 0.84	4.04 ± 1.57	4.13 ± 2.00	3.73 ± 1.77
	TT (n=41)	183.1 ± 35.8	174.10 ± 38.10	181.10 ± 35.52	1.49 ± 0.87	1.98 ± 1.21	1.40 ± 0.78	4.15 ± 1.61	5.07 ± 2.62	3.71 ± 1.60
Jung (24) / 2013	Subplatysmal (n=42)	169.7 ± 47.9	169.80 ± 43.40	170.40 ± 46.10	1.01 ± 0.79	0.99 ± 1.05	0.90 ± 0.79	3.38 ± 2.07	3.35 ± 1.88	3.34 ± 2.05
	Subfascial (n=44)	180.4 ± 30.5	180.00 ± 31.30	184.50 ± 31.80	1.08 ± 0.76	0.98 ± 1.04	1.53 ± 3.02	3.58 ± 2.12	3.67 ± 2.02	3.70 ± 2.38
Overall mean		**192.7 ± 35.6**	**188.1 ± 34.8***	**190.5 ± 38.1#**	**0.90 ± 0.62**	**0.99 ± 0.72#**	**0.92 ± 1.07#**	**3.27 ± 1.58**	**3.24 ± 1.71***	**3.12 ± 1.73#**

Abbreviations: <3-m = less-than 3 months of thyroidectomy; ≥3-m = 3 months or more after thyroidectomy; LTT = less-than total thyroidectomy; TT = total thyroidectomy

* Significant (p<0.05) relative to baseline by fixed or random-effect model

Not significant (p≥0.05) relative to baseline by fixed or random-effect model

Table 3. Relative changes in noise-to-harmonic ratio and maximum phonation time < 3 months and ≥3 months after thyroidectomy.

First author / year	Subgroups available	Noise-to-Harmonic ratio (NHR)			Maximum phonation time (MPT) (seconds)		
		Baseline	**<3m**	**≥3m**	**Baseline**	**<3m**	**≥3m**
Hong (5) / 1997	No	-	-	-	21.7 ± 7.0	18.80 ± 7.60	19.65 ± 8.05
Debruyne (11) / 1997	No	-	-	-	-	-	-
Aluffi (12) / 2001	No	0.12	-	0.13	-	-	-
Stojadinovic (3) / 2002	LTT (n=21)	0.12 ± 0.04	0.13 ± 0.03	0.12 ± 0.04	18.08 ± 4.33	19.05 ± 4.66	18.42 ± 6.18
	TT (n=29)	0.12 ± 0.03	0.13 ± 0.05	0.13 ± 0.05	17.35 ± 5.57	17.17 ± 4.78	18.36 ± 6.81
Sinagra (13) / 2004	No	-	-	-	-	-	-
Netto (14) / 2006	Males (n=11)	0.10 ± 0.02	0.20 ± 0.03	-	-	-	-
	Females (n=77)	0.20 ± 0.03	0.10 ± 0.02	-	-	-	-
Soylu (15) / 2007	LTT (n=8)	0.121 ± 0.06	0.13 ± 0.09	0.12 ± 0.07	-	-	-
	TT (n=40)	0.113 ± 0.03	0.14 ± 0.03	0.12 ± 0.02	-	-	-
Akyildiz (16) / 2008	Males (n=9)	0.13 ± 0.01	0.14 ± 0.02	-	-	-	-
	Females (n=27)	0.15 ± 0.04	0.14 ± 0.03	-	-	-	-
Lombardi (17) / 2009	No	0.10 ± 0.01	-	0.10 ± 0.02	20.2 ± 6.50	-	19.80 ± 5.80
Van Lierde (18) / 2010	Males (n=12)	-	-	-	25.3 ± 9.89	23.40 ± 5.45	30.20 ± 13.10
	Females (n=32)	-	-	-	18.5 ± 4.72	18.28 ± 5.00	18.27 ± 6.49
Li (19) / 2012	Males (n=8)	0.13 ± 0.02	0.14 ± 0.02	0.12 ± 0.01	-	-	-
	Females (n=24)	0.15 ± 0.04	0.14 ± 0.05	0.10 ± 0.02	-	-	-
Lombardi (20) / 2012	No	0.11 ± 0.01	-	0.14 ± 0.01	17.8 ± 6.60	-	19.30 ± 6.10
Tae (21) / 2012	No	0.14 ± 0.04	0.22 ± 0.83	0.13 ± 0.04	12.91 ± 3.91	11.63 ± 4.16	12.43 ± 4.22
Lee (22) / 2012	No	0.13 ± 0.04	0.12 ± 0.06	0.13 ± 0.04	11.5 ± 4.0	10.00 ± 4.00	11.60 ± 4.00
Ryu (23) / 2013	No	-	-	-	-	-	-
Jung (24) / 2013	Subplatysmal (n=42)	0.13 ± 0.03	0.13 ± 0.02	0.12 ± 0.02	-	-	-
	Subfascial (n=44)	0.13 ± 0.03	0.13 ± 0.03	0.13 ± 0.03	-	-	-
Overall mean		**0.13 ± 0.03**	**0.14 ± 0.31#**	**0.12 ± 0.03#**	**17.8 ± 5.8**	**15.5 ± 5.3***	**17.7 ± 6.3#**

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

* Significant ($p<0.05$) relative to baseline by fixed or random-effect model

Not significant ($p\geq0.05$) relative to baseline by fixed or random-effect model

Table 4a. A comparison of F_0 changes in early period between male and female subgroups*

Male						Female				
		Baseline		**<3 months**			**Baseline**		**<3 months**	
	n	Mean	SD	Mean	SD	n	Mean	SD	Mean	SD
Netto (14) / 2006	11	127.65	13.60	117.43	16.70	77	236.81	52.50	231.80	44.70
Akyildiz (16) / 2008	9	114.20	23.00	110.30	7.50	27	196.80	20.60	197.80	27.10
Van Lierde (18) / 2010	12	129.26	17.90	118.06	28.20	32	220.43	31.10	215.80	24.70
Li (19) / 2012	8	115.30	18.10	104.75	17.60	24	191.70	25.40	189.10	27.30
		Estimate	**Lower limit**	**Upper limit**	***p*-value**		**Estimate**	**Lower limit**	**Upper limit**	***p*-value**
Fixed-effect estimate		-0.470	-0.935	-0.005	**0.048**		-0.043	-0.262	0.177	0.704
Random-effect estimate		-0.470	-0.935	-0.005	**0.048**		-0.043	-0.262	0.177	0.704

* No significant differences in F_0 were found between baseline and late (≥3 months) period for both males and females (*data not shown*)

Table 4b. A comparison of F_0 changes in early period between after total thyroidectomy and less-than total thyroidectomy

Total thyroidectomy						Less-than total thyroidectomy				
		Baseline		**<3 months**			**Baseline**		**<3 months**	
	n	Mean	SD	Mean	SD	n	Mean	SD	Mean	SD
Debruyne (11) / 1997	47	210.40	24.00	205.40	19.70	29	165.60	40.70	168.07	45.50
Stojadinovic (3) / 2002	21	169.23	37.90	170.82	44.40	46	211.00	17.00	176.00	19.00
Soylu (15) / 2007	8	236.54	69.30	217.88	71.80	40	237.71	32.00	221.64	26.30
Li (19) / 2012	32	209.52	46.44	203.21	39.96	-	-	-	-	-
Ryu (23) / 2013	33	189.60	44.00	187.70	50.50	41	183.10	35.80	174.10	38.10
		Estimate	**Lower limit**	**Upper limit**	***p*-value**		**Estimate**	**Lower limit**	**Upper limit**	***p*-value**
Fixed-effect estimate		-0.486	-0.740	-0.232	**<0.001**		-0.127	-0.361	0.108	0.291
Random-effect overall estimate		-0.633	-1.339	-0.074	**0.047**		-0.127	-0.361	0.108	0.291

LEGENDS

Appendix 1. A list of articles which were reviewed but not included in the study

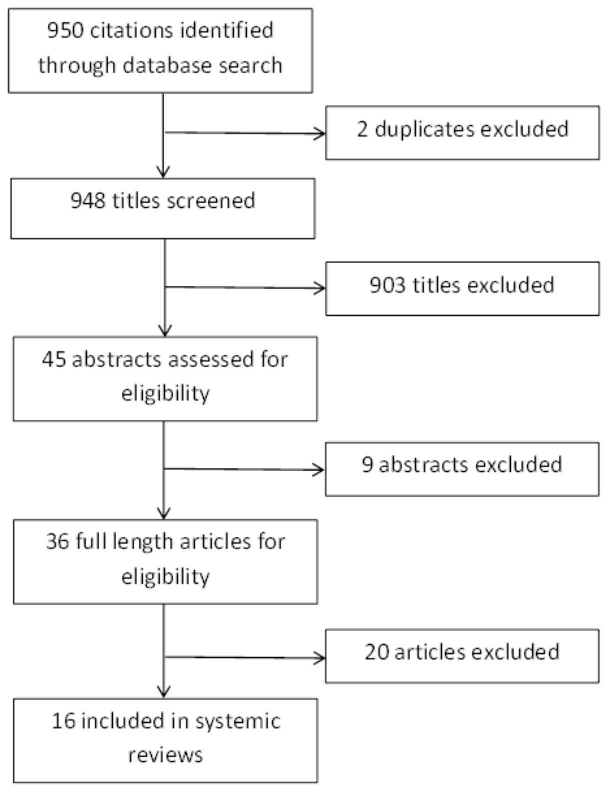
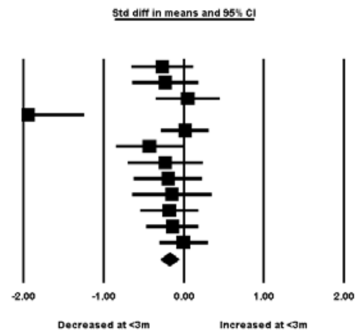


Figure 1. A flow diagram for study selection

Model	Study name	Total	Pre-op Mean	Pre-op SD	<3m Mean	<3m SD	Statistics for each study			
							Std diff in means	Lower limit	Upper limit	p-Value
	Hong / 1997	54	207.000	47.000	196.000	35.000	-0.265	-0.649	0.118	0.175
	Debruyne / 1997	47	210.400	24.000	205.400	19.700	-0.228	-0.637	0.182	0.276
	Stojadinovic / 2002	50	167.125	39.557	169.225	45.045	0.050	-0.343	0.442	0.804
	Sinagra / 2004	46	211.000	17.000	176.000	19.000	-1.941	-2.636	-1.247	0.000
	Netto / 2006	88	186.475	20.893	186.863	25.604	0.017	-0.279	0.312	0.912
	Soylu / 2007	48	237.515	39.987	221.013	37.026	-0.428	-0.846	-0.010	0.045
	Akyildiz / 2008	36	197.638	28.549	191.365	25.567	-0.231	-0.700	0.237	0.332
	Van Lierde M / 2010	44	170.864	23.706	166.095	25.124	-0.195	-0.617	0.227	0.364
	Li / 2012	32	209.520	46.436	203.208	39.962	-0.146	-0.638	0.347	0.562
	Tae / 2012	61	207.900	37.000	201.400	34.900	-0.181	-0.539	0.177	0.322
	Ryu / 2013	74	185.999	39.654	180.165	44.044	-0.139	-0.463	0.185	0.399
	Jung / 2013	86	175.174	39.951	175.019	37.694	-0.004	-0.303	0.295	0.979
Fixed							-0.180	-0.291	-0.069	0.001



Model	Study name	Total	Pre-op Mean	Pre-op SD	>=3m Mean	>=3m SD	Statistics for each study			
							Std diff in means	Lower limit	Upper limit	p-Value
	Hong / 1997	54	207.000	47.000	209.000	45.400	0.043	-0.334	0.421	0.822
	Stojadinovic / 2002	50	167.125	39.557	167.984	40.786	0.021	-0.371	0.413	0.915
	Sinagra / 2004	46	211.000	17.000	192.000	13.500	-1.238	-1.781	-0.695	0.000
	Soylu / 2007	48	237.515	39.987	215.882	39.860	-0.542	-0.970	-0.113	0.013
	Lombardi / 2009	110	186.800	38.700	185.000	36.300	-0.048	-0.312	0.216	0.722
	Van Lierde M / 2010	44	170.864	23.706	169.618	24.903	-0.051	-0.469	0.367	0.810
	Li / 2012	32	209.520	46.436	213.340	34.470	0.093	-0.398	0.584	0.709
	Lombardi / 2012	32	175.000	30.000	179.300	32.700	0.137	-0.355	0.629	0.585
	Tae / 2012	61	207.900	37.000	202.150	32.800	-0.164	-0.522	0.193	0.367
	Ryu / 2013	74	185.999	39.654	181.256	42.925	-0.115	-0.438	0.209	0.487
	Jung / 2013	86	175.174	39.951	177.614	39.433	0.061	-0.238	0.361	0.687
Fixed							-0.106	-0.219	0.006	0.063

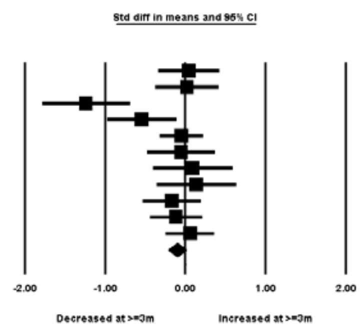
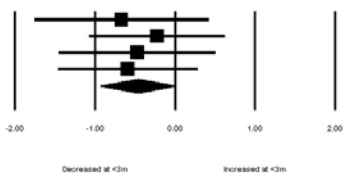


Figure 2. Forest plots showing the standardized mean difference in fundamental frequency (F_0, Hz) between preoperative and <3 months and ≥3 months of thyroidectomy

Model	Study name	Total	Pre-op Mean	Pre-op SD	<3m Mean	<3m SD	Std diff in means	Lower limit	Upper limit	p-Value
	Li / 2012	8	127.660	13.600	117.430	16.700	-0.671	-1.756	0.414	0.226
	Netto / 2006	11	114.200	23.000	110.300	7.500	-0.228	-1.075	0.619	0.598
	Akyildiz / 2008	9	129.260	17.900	118.060	28.200	-0.474	-1.449	0.500	0.340
	Van Lierde / 2010	12	115.300	10.100	104.750	17.600	-0.591	-1.458	0.276	0.182
Fixed							-0.470	-0.935	-0.005	0.046



Model	Study name	Pre-op Mean	Pre-op SD	<3m Mean	<3m SD	Std diff in means	Lower limit	Upper limit	p-Value
	Li / 2012	236.810	52.500	231.800	44.700	-0.103	-0.670	0.465	0.723
	Netto / 2006	186.800	20.600	197.800	27.100	0.042	-0.274	0.358	0.797
	Akyildiz / 2008	220.430	31.100	215.800	24.700	-0.165	-0.702	0.372	0.547
	Van Lierde / 2010	191.700	25.400	189.100	27.300	-0.099	-0.590	0.393	0.694
Fixed						-0.043	-0.262	0.177	0.704

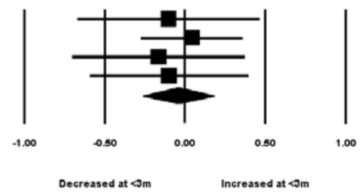


Figure 3. Forest plots showing the standardized mean difference in fundamental frequency (F_0, Hz) between preoperative and <3 months of thyroidectomy in males and females

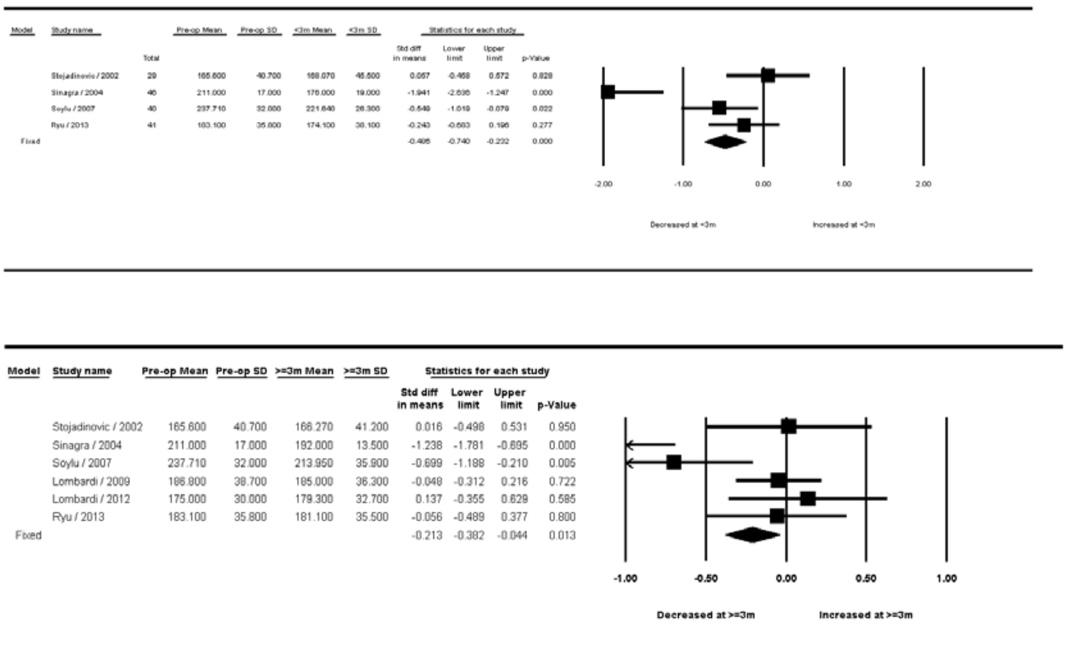


Figure 4. Forest plots showing the standardized mean difference in fundamental frequency (F_0, Hz) between preoperative and <3 months and ≥3 months after total thyroidectmy.