Quantitative study of voice dysfunction after thyroidectomy

James C. Lee, MBBS, FRACS, PhD,a,b Daniel Breen, MBBS,a Amanda Scott, PhD,c Simon Grodski, MBBS, FRACS,a,b Meei Yeung, MBBS, BMedSci, FRACS,a William Johnson, MBBS, MD, FRACS, FACS,a and Jonathan Serpell, MBBS, MD, MEd, FRACS, FACS,a,b Prahran, Victoria, Australia

Background. Up to 80% of patients without a recurrent laryngeal nerve palsy report alteration in their voice after a thyroid procedure. The aims of this study were (1) to quantify voice changes after thyroid operation; (2) to correlate the changes to the extent of operation; and (3) to correlate voice changes to intraoperative recurrent laryngeal nerve swelling.

Methods. Patients undergoing total and hemithyroidectomy were recruited prospectively from the Monash University Endocrine Surgery Unit during a 12-month period. Voice quality was scored subjectively using the Voice Disorder Index (score 0–40, from best to worst) and objectively using the Dysphonia Severity Index (score 0–5, from worst to best), before and after operation. These assessments were carried out by 2 speech pathologists. Recurrent laryngeal nerve diameter was measured intraoperatively at the commencement and conclusion of the lobectomy, using Vernier calipers with a resolution of 0.1 mm. Statistical methods used included Student t test, χ², Wilcoxon signed-rank test, and linear regression.

Results. A total of 62 patients were included in the study, with a mean age of 48 ± 16 years and a female preponderance (6:1). Overall, the voice quality deteriorated both subjectively (mean Voice Disorder Index 4.2 ± 0.8–9.4 ± 1.2, P < .01) and objectively (mean Dysphonia Severity Index 3.9 ± 0.3–2.8 ± 0.3, P < .01) with thyroid operation. Patients who underwent either hemi- or total thyroidectomy both reported significant deterioration of voice (mean Voice Disorder Index 5.4 ± 1.5–7.9 ± 1.4, P = .02 and 3.4 ± 0.7–10.4 ± 1.8, P < .01 respectively). However, on objective assessment, only total thyroidectomy patients showed significant deterioration (Mean Dysphonia Severity Index 4.0 ± 0.3–2.5 ± 0.3, P < .01). At 6–12 months, both Voice Disorder Index and DSI returned to preoperative levels. Intraoperatively, the recurrent laryngeal nerve diameter increased by 0.58 ± 0.05 mm (1.82 ± 0.05 mm–2.40 ± 0.05 mm; P < .01). In hemithyroidectomy patients, the degree of nerve swelling correlated with the degree of deterioration in objective voice assessment, in that the greater the increase in recurrent laryngeal nerve diameter, the worse the Dysphonia Severity Index score (coefficient −0.4, P = .03). This was not the case in the total thyroidectomy patients.

Conclusion. Voice quality deteriorates with thyroid operation despite functionally intact recurrent laryngeal nerve. While likely multifactorial, the degree of deterioration is related to the extent of operation and may also be related to the degree of recurrent laryngeal nerve swelling. Spontaneous resolution is expected in the majority of patients. (Surgery 2016;160:1576-81.)
The human voice is produced by a complex instrument comprising of intricate laryngeal structures, which are in turn innervated by branches of the recurrent laryngeal nerve, aided by branches of the superior laryngeal nerve. The effect on the voice of a RLN palsy is recognized and documented easily. However, the subtler changes to the voice in the absence of a nerve palsy are less well documented.

With the overarching aim of better understanding the potential causes of voice change during thyroid operation, the aims of this study were to quantify the voice change after thyroid operation; to correlate the measured voice change to the extent of operation; and to correlate the degree of voice change to the amount of recurrent laryngeal nerve swelling.

METHODS

Ethics approval was granted by the institutional review board prior to the commencement of this study. The study patients were prospectively recruited during a 12-month period. All patients undergoing total and hemithyroidectomy in the authors’ institution were eligible. However, patients having a concurrent lymph node dissection were excluded. All participating patients had preoperative and postoperative flexible laryngoscopy to assess vocal cord function and underwent functional voice assessments using established tools. The voice was assessed both subjectively and objectively by speech pathologists using a professional speech and voice analysis system (WEVOSYS lingWAVES 2.5, WEVOSYS, Forchheim, Germany).

The Voice Disorder Index (VDI) was used for the subjective assessment. An abbreviated version of the Voice Handicap Index, it is a self-reported measure of suffering caused by dysphonia. This instrument gives insight into the voice-related quality of life by quantifying the voice problem as it relates to the common activities in daily living. The VDI score ranges from 0–40, from best to worst. It has been demonstrated that this abbreviated assessment performed equally well compared with the extended counterpart.

The Dysphonia Severity Index (DSI) was used for the objective assessment. It was designed to establish an objective and quantitative correlate of the perceived voice quality. The DSI is based on the weighted combination of highest frequency, lowest intensity, maximum phonation time, and jitter. A perceptually normal voice would score a DSI of equal to or greater than +5 and a severely dysphonic voice would have a DSI of −5 or less, with +5 to −5 being the usual range. The DSI has been used to objectively monitor the result of voice therapy and training programs and has been shown to have good interobserver consistency. Although there is no consensus on how the voice is best assessed in order to determine the effect of thyroidectomy, these 2 assessment tools were chosen as they have been validated previously and are relatively simple to administer, without any invasive procedure.

The preoperative assessments were performed up to 6 weeks preoperatively, and the postoperative assessments were performed day 1 postoperatively prior to discharge. Follow-up assessments were performed between 6–12 months postoperatively.

Thyroidectomy was performed as previously described. Dexamethasone (8 mg) was administered routinely at the beginning of each case. Intraoperatively, the RLN diameter was measured at 2 distinct and specific time points. It was first measured when the RLN was first identified and confirmed by neuromonitoring. The second measurement was taken immediately upon completion of the lobectomy. These measurements were made using Vernier calipers with a resolution of 0.1 mm. Intraoperative neuromonitoring (Medtronic NIM 3.0, Medtronic, Minneapolis, MN) is routinely used in all cases.

Standard statistical methods were used to analyze the collected data, including t test, \( \chi^2 \) test, Wilcoxon signed-rank test, Pearson correlation, and linear regression. Stata version 12 was used for statistical analysis.

RESULTS

Of the 70 recruited patients, 62 (89%) completed all the necessary assessments and were included in the analysis. The mean age of the study cohort was 47.7 ± 16 years (range, 18–80 years). There were 53 females and 9 males, giving a ratio of 6:1. Thirty-seven (60%) patients had a total thyroidectomy, while 25 (40%) had a hemithyroidectomy. Temporary RLN palsy occurred in 4 patients (6.5%), and there was no permanent RLN palsy (Table I).

In the overall cohort, the VDI score (subjective assessment) increased from 4.2 preoperatively to 9.4 postoperatively (\( P < .01 \)). This subjective deterioration in the voice also was seen in patients without postoperative RLN palsy, not just in the 4 patients who sustained a palsy (Table II). On average, the postoperative VDI score of the patients with palsy was 19.9 points higher than those without a palsy (\( P < .01 \)); interestingly, patients...
with a postoperative palsy also had a higher VDI score than the nonpalsy patients ($P = .02$) before operation (Table II).

The DSI scores (objective assessment) comparing preoperative and postoperative voice assessments showed a trend very similar to that of the VDI scores. The overall score deteriorated from 3.9 to 2.8 ($P < .01$). Significant deterioration in DSI also was seen in both the palsy and nonpalsy subgroups ($P = .042$ and $P < .01$ respectively; Table III). As expected, patients with a postoperative palsy had worse DSI scores ($P < .01$), but once again, they also had worse DSI scores before undergoing operation ($P = .03$).

Patients having either a total thyroidectomy or a hemithyroidectomy both showed significant deterioration in their VDI scores ($P < .01$ and .02, respectively; Tables IV and V). However, objectively, the DSI scores deterioration postoperatively was only seen in patients undergoing a total thyroidectomy ($P < .01$), not those undergoing a hemithyroidectomy ($P = .1$; Tables IV and V).

A small group of 13 patients attended the follow-up assessments, including 3 patients with a temporary unilateral RLN palsy (Table VI). Of these, 6 patients had hemithyroidectomy and 7 had total thyroidectomy. The median follow-up was 8 months (range, 6–14 months). In this group of patients, the follow-up mean VDI was 4.2 ± 0.8, no different from their mean preoperative VDI of 7.1 ± 2.7 ($P = .26$). Their follow-up mean DSI was 5.4 ± 0.9, even better their mean preoperative DSI of 3.5 ± 0.9 ($P = .01$).

Collectively, the RLN diameter increased by 0.58 ± 0.05 mm (1.82 ± 0.05 mm to 2.40 ± 0.05 mm; $P < .01$). Using Pearson correlation, it was demonstrated that the DSI scores of hemithyroidectomy patients correlated to the degree of RLN swelling observed intraoperatively. This correlation was not shown in total thyroidectomy patients or the DSI scores of hemithyroidectomy patients (Table VII, Fig).

**DISCUSSION**

The data in this study suggest that the quality of voice deteriorated after thyroid operation, both subjectively and objectively. In addition, this deterioration was more pronounced in total thyroidectomy patients than hemithyroidectomy patients.
Spontaneous recovery of the voice was seen when patients returned for long-term follow-up. Furthermore, the objective voice deterioration correlated to the degree of RLN swelling in hemithyroidectomy patients.

An interesting finding in this study was that although the extent of operation did not correlate to the degree of subjective voice deterioration, only total thyroidectomy patients were found to have a significant deterioration in their voice on objective assessment. In other words, on objective assessment with DSI, the degree of voice deterioration correlated to the extent of operation in our patients. We think that a RLN that has been manipulated during operation do not function optimally for a period of time postoperatively, due to minor damage that is not enough to cause a palsy. In hemithyroidectomy patients, the vocal cord on the nonoperated side is able to compensate the subtle deterioration of the vocal cord of operated side. In total thyroidectomy patients, both vocal cords are functioning suboptimally, and thus scored more poorly in the DSI than hemithyroidectomy patients.

This finding is in contrast to a 2002 Memorial Sloan Kettering Cancer Center (MSKCC) study by Stojadinovic et al on a cohort of 54 patients undergoing thyroidectomy (partial, subtotal, or total). The authors from MSKCC reported that the extent of thyroidectomy did not correlate significantly with the degree of subjective or objective voice changes. In their study, Stojadinovic et al utilized acoustic, aerodynamic, glottographic, and videostroboscopic tests to assess the voice objectively. The different voice assessment instruments are a potential reason for the differing conclusions in these studies.

The most novel finding in this study is that the degree of objective voice deterioration correlated to the degree of RLN swelling seen intraoperatively in hemithyroidectomy patients. Although RLN swelling during the course of thyroidectomy is a phenomenon that has been reported previously, this is the first time that the swelling has been shown to correlate to functional outcome. If the degree of RLN swelling is a surrogate for varying degree of nerve damage without reaching the stage of palsy, then it would make sense that the degree of voice deterioration correlated with it. The causal effect of this correlation, and whether there are ways to minimize RLN swelling in order to minimize voice change, remain topics of ongoing investigations.

The findings of this study support the notion that voice change post thyroidectomy is a complex phenomenon, more than the result of RLN palsy, as indicated by other authors. Apart from partial and temporary RLN damage, other potential factors affecting post-thyroidectomy voice changes include injuries to the extrinsic laryngeal muscles, patient’s smoking history, the effect of endotracheal intubation, and injury to the external branch of the superior laryngeal nerve. One of the limitations of this study is that effects of these factors were not specifically sought.

### Table IV. Voice quality change with different extent of operation for the Voice Disorder Index (0 best, 40 worst)

<table>
<thead>
<tr>
<th></th>
<th>Hemithyroidectomy (N = 25)</th>
<th>Total thyroidectomy (N = 37)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean preoperative score (SE)</td>
<td>5.4 (1.5)</td>
<td>3.4 (0.7)</td>
</tr>
<tr>
<td>Mean postoperative score (SE)</td>
<td>7.9 (1.4)</td>
<td>10.4 (1.8)</td>
</tr>
<tr>
<td>VDI change (SE)</td>
<td>2.5 (1.0)</td>
<td>6.9 (1.9)</td>
</tr>
<tr>
<td>P value</td>
<td>.02</td>
<td>&lt;.01</td>
</tr>
</tbody>
</table>

VDI, Voice Disorder Index.

### Table V. Voice quality change with different extent of operation for the Dysphonia Severity Index (worst −5, best 5)

<table>
<thead>
<tr>
<th></th>
<th>Hemithyroidectomy (N = 25)</th>
<th>Total thyroidectomy (N = 37)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean preoperative score (SE)</td>
<td>3.8 (0.4)</td>
<td>4.0 (0.3)</td>
</tr>
<tr>
<td>Mean postoperative score (SE)</td>
<td>3.1 (0.4)</td>
<td>2.5 (0.3)</td>
</tr>
<tr>
<td>DSI change</td>
<td>−0.66 (0.4)</td>
<td>−1.5 (0.3)</td>
</tr>
<tr>
<td>P value</td>
<td>.1</td>
<td>&lt;.01</td>
</tr>
</tbody>
</table>

DSI, Dysphonia Severity Index; SE, standard error.

### Table VI. Follow-up (median, 8 months) for Voice Disorder Index and Dysphonia Severity Index

<table>
<thead>
<tr>
<th></th>
<th>Voice Disorder Index (N = 13)</th>
<th>Dysphonia Severity Index (N = 13)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean preoperative score (SE)</td>
<td>7.1 (2.7)</td>
<td>3.5 (0.9)</td>
</tr>
<tr>
<td>Mean follow-up score (SE)</td>
<td>4.2 (0.8)</td>
<td>5.4 (0.9)</td>
</tr>
<tr>
<td>P value</td>
<td>.26</td>
<td>.01</td>
</tr>
</tbody>
</table>

SE, Standard error.
However, as the focus was to compare patients undergoing different extent of thyroid operation, the effect of these factors play a lesser role in the comparisons.

With only 20% completion rate of the follow-up assessments, it is difficult to determine if the available follow-up data can be extrapolated to the entire study population. Nonetheless, it was still reassuring to find that the mean VDI score had returned to a level similar to that preoperatively and that the DSI score was even greater than the preoperative level (Table V). These findings confirm what others have found: Post-thyroidectomy voice deterioration resolves spontaneously. In a 2010 study, no permanent (>3 months) change to the vocal performance was found in patients after thyroidectomy without RLN injury. It has been reported that it takes between 2 weeks and a few months for the majority of voice changes to resolve.1,15,16

Several limitations of this study need to be recognized. The study design did not include a nonthyroid operation group; therefore, the effect of endotracheal intubation cannot be studied. Similarly, other potential factors affecting postoperative voice changes were beyond the scope of this study. Specifically, this study is unable to comment on the external branch of the superior laryngeal nerve in the study patients. Although not strictly a limitation, the use of different assessment instruments may make the findings in this study not entirely comparable with other studies. Finally, the high attrition rate at follow-up beyond 6 months makes extrapolation of the follow-up data difficult.

It is suggested that voice changes in the early postoperative period may be useful to predict patients who would benefit from voice therapy, either to achieve more rapid or more complete resolution.2,17 Therefore, this study serves as a foundation for future studies. With a larger cohort and more stringent follow-up protocol, it is hoped that an extension of this study will allow us to investigate whether a combination of VDI and DSI assessments can be used to select patients who would most benefit from voice therapy.

In conclusion, there is no doubt that the voice quality can deteriorate after thyroidectomy and that patients undergoing total thyroidectomy can expect more deterioration than patients undergoing hemithyroidectomy. However, it is reassuring to know that the majority of cases recover spontaneously by 6–12 months. This information is very useful in preoperative patient counselling and managing expectations.

**REFERENCES**

