The Impact of Post-thyroidectomy Paresis on Quality of Life in Patients with Nodular Thyroid Disease

Jesper Roed Sorensen, PhD1, Trine Printz, MA1, Jenny Iwarsson, PhD2, Ågot Møller Grøntved, MD1, Helle Døssing, PhD1, Laszlo Hegeduš3, Steen Joop Bonnema3, Christian Godballe1, and Camilla Slot Mehlum, MD1

Abstract

Objective. To investigate the impact of postoperative paresis on disease-specific quality of life (DSQoL) after thyroidectomy in patients with benign nodular thyroid disease.

Study Design. Observational study.

Setting. University hospital.

Subjects and Methods. Patients were evaluated before and 3 weeks and 6 months after surgery in an individual prospective cohort study using videolaryngostroboscopy (VLS), voice range profile, voice handicap index (VHI), multidimensional voice program, maximum phonation time (MPT), and auditory perceptual evaluation. Changes in DSQoL were assessed by the Thyroid-specific Patient-Reported Outcome measure. Cohen’s effect size was used to evaluate changes.

Results. Sixty-two patients were included, 55 of whom completed all examinations. Three weeks after surgery, a blinded VLS examination showed signs of paresis of either the recurrent laryngeal nerve or the external branch of the superior laryngeal nerve (RLN/EBSLN) in 13 patients (24%). A paresis corresponded to a 12±28 point increase in VHI (P=.002) and was associated with a significant 4.3±6.7 semitone decrease in the maximum fundamental frequency (P<.001) and a 5.3±8.2 dB reduction in maximum intensity. Further, it was associated with a 4.5±11.2 second reduction in MPT (P=.001) and an increase of 0.40±1.19 in grade, 0.42±1.41 in roughness, and 0.36±1.11 in breathiness. Signs of postoperative RLN/EBSLN paresis correlated with an 11.0-point (P=.02) poorer improvement in goiter symptoms at both 3 weeks and 6 months after surgery.

Conclusion. Signs of RLN/EBSLN paresis after thyroidectomy were associated with less pronounced improvement in goiter symptoms in patients with thyroid nodular disease. However, thyroidectomy was associated with an overall improved DSQoL by 6 months after surgery.

Keywords

thyroidectomy, goiter, recurrent laryngeal nerve, external branch of the superior laryngeal nerve, voice, quality of life

Received January 7, 2019; accepted May 16, 2019.

 Thyroidectomy is a well-established treatment option for patients with thyroid nodularity, Graves’ disease, and thyroid malignancy.1-4 Thyroid surgery profoundly improves disease-specific quality of life (DSQoL) for patients with benign as well as malignant thyroid disease5-9 and relieves both tracheal and esophageal compression.10,11 When performed at high-volume centers, thyroidectomy is reported to be associated with a low rate of long-term complications (ie, hypocalcemia and recurrent laryngeal nerve [RLN] paralysis).12-15 However, up to 40% of patients experience a postoperative voice, vocal fold change, or both,16-20 which can be caused by intubation injuries,21 paresis of the external branch of the superior laryngeal nerve (EBSLN),22,23 RLN paresis,20 or maladaptive mechanisms.16 Especially RLN or EBSLN (RLN/EBSLN) paresis can cause permanent voice changes.20,22,23 Most postoperative voice, vocal fold changes, or both disappear within 3 to 6 months after surgery, but the frequency of paresis might be grossly underestimated.16,18,24,26 Previous studies have indicated a detrimental impact of voice disorders on...
patients’ quality of life (QoL). Knowledge of how RLN/EBSLN paresis affects symptom improvement based on DSQoL in patients with thyroid nodularity is sparse yet pertinent to obtain, as nonsurgical treatment options have increasingly been implemented.4,28-35

We aimed at investigating the association between RLN/EBSLN paresis and symptom improvement, based on DSQoL, in patients with benign thyroid nodular disease undergoing thyroidectomy.

Materials and Methods

Patient Characteristics

In a tertiary referral setting, patients with thyroid nodularity scheduled for thyroid lobectomy or total thyroidectomy were consecutively included in a study with 6 months of postoperative follow-up. Thyroid nodularity was defined as the presence of an enlarged thyroid gland on ultrasound with at least 1 nodule confirmed by ultrasonography. Serum levels of total thyroxine (T4) and total triiodothyronine (T3) had to be within the normal range (67-134 nmol/L and 1.35-2.33 nmol/L, respectively), but subclinically hypo- or hyperthyroid patients were eligible. Patients were included from November 2014 with the last visits in August 2017, after a follow-up period of 6 months. The exclusion criteria were (1) previous surgery to the neck, including thyroid surgery; (2) suspicion of thyroid cancer, as these patients were managed in these patients.42 This study was part of a larger study investigating changes in QoL, from which some of the data were previously published but in a different context.9,10,43,44

Vocal Fold and Voice Outcome Measures

Each patient underwent a preoperative multifactorial voice and vocal fold examination, which was repeated 3 weeks and 6 months after surgery. It was performed by an experienced resident in ear, nose, and throat diseases. This program comprised a videolaryngostroboscopic (VLS) examination conducted with a 70° rigid videolaryngoscope (Pulsar II, Karl Storz, Germany) with a stroboscopic light source. Two experienced laryngologists independently rated the blinded videos after completion of the study. The consultants addressed vocal fold mobility, vocal fold lengthening during intonation, and potential localized vocal fold lesion as dichotomized variables. Impaired vocal fold mobility was interpreted as a RLN paresis based on normal mobility preoperatively and the absence of long-term intubation or other disease that might cause immobility by fixation. Impaired ability to lengthen the vocal folds was interpreted as “sign of EBSLN paresis,” because postoperative discomfort and maladaptive mechanisms might mimic EBSLN palsy. The consultants had no access to information on how well the patient had cooperated at the time of examination. Any discrepancies were resolved by discussion until consensus was reached.

The dual-microphone phonetogram system Voice Profiler 5.0 (Alphatron Medical Systems, Rotterdam, the Netherlands) was used for assessing the voice range profile.37 The parameters of interest were the minimum and maximum fundamental frequencies (f0), measured in semitones (ST); the frequency range; the minimum and maximum voice intensity, defined as sound pressure level (SPL), measured in decibels (dB); and the intensity range.

Jitter (%), shimmer (%), noise to harmonic ratio (NHR), voice turbulence index (VTI), and soft phonation index (SPI) were assessed by the multidimensional voice program (MDVP; model 4400, KeyPENTAX, Lincoln Park, New Jersey, USA). Patients produced the sustained vowel /a:/ at a comfortable frequency and intensity. The procedure was repeated 3 times, and the median value was used for analyses.

The maximum phonation time (MPT) was recorded having the patient sustain the vowel /a:/ for as long as possible on a single breath using a handheld recorder (LS-11, Olympus, Tokyo, Japan). This was repeated 3 times, and the longest sample was used for analyses.

A blinded auditory-perceptual evaluation was performed by a speech-language pathologist using the most reliable parameters of the Grade, Roughness, Breathiness, Asthenia, Strain (GRBAS) scale system.38 The voice samples were 15 seconds of reading the Danish version of the “North Wind and the Sun.” Each voice sample was rated on the G, R, and B scales using a 4-point Likert-type scale (0-3), with higher scores indicating greater abnormality of the voice.

The voice handicap index (VHI) questionnaire was completed prior to the VLS examination. This well-validated questionnaire39,40 comprises 30 statements regarding patient-reported severity of voice handicap. Each statement is scored on a 5-point Likert-type scale ranging from 0 (never) to 4 (always), with a possible score of 120 points. Thus, a higher VHI score indicates a greater self-reported impact or “handicap” of a voice problem. A validation study of the Danish VHI found control persons having a VHI score of 6.5 ± 8.3 points.40

Symptom Improvement Based on DSQoL

Symptom improvement, based on DSQoL, was assessed before surgery and 3 weeks and 6 months after surgery using the Thyroid-specific Patient-Reported Outcome measure (ThyPRO). This is a self-administered questionnaire41 with 13 multi-item scales covering symptom improvement after thyroidectomy (Goiter Symptoms, Hyperthyroid Symptoms, Hypothyroid Symptoms, Eye Symptoms, Tiredness, Cognitive Complaints, Anxiety, Depressivity, Emotional Susceptibility, Impaired Social Life, Impaired Daily Life, Impaired Sex Life, and Cosmetic Complaints), as well as a single item measuring the overall impact of thyroid disease on QoL. The questionnaire includes 85 items rated on a 5-point Likert-type scale ranging from 0 (not at all) to 4 (very much). Each scale was scored as a summary score and linearly transformed to a range of 0 to 100 points, with a lower score indicating improved health status. The Eye Symptom scale was omitted as it was not relevant for these patients.42 This study was part of a larger study investigating changes in QoL, from which some of the data were previously published but in a different context.9,10,43,44
Surgical Procedure

The thyroid surgery comprised either a thyroid lobectomy or a total extracapsular thyroidectomy and was performed under general anesthesia by a consultant head and neck surgeon. The RLN was monitored using the NIM EMG tube (Medtronic, Minneapolis, Minnesota, USA) and was exposed to the inferior constrictor. Patients were discharged the first postoperative day in case of uncomplicated lobectomy. After total thyroidectomy, levothyroxine was administered. These patients were discharged when calcium levels approached the normal range. Patients were followed up at 2 to 4 weeks after surgery and again 6 to 10 weeks after surgery.

Statistics

A 3-step statistical approach was used for assessing the impact of RLN/EBSLN paresis on DSQoL. Initially, a linear mixed model was used to evaluate the impact of paresis on changes in voice parameters from the multifactorial voice examination from baseline to 3 weeks and 6 months after surgery. The independent variables were age, sex, thyroid volume, surgical procedure, RLN/EBSLN paresis, and time point (baseline, 3 weeks after surgery, and 6 months after surgery). The linear mixed model was used to interpret the significance of the changed objective voice parameters (parameters with significant and moderate-sized changes) on VHI score using age, sex, thyroid volume, surgical procedure, time, RLN/EBSLN paresis, maximum $f_0$, and maximum SPL as independent variables. A linear mixed model was used to evaluate the association of RLN/EBSLN paresis on changes in ThyPRO scale scores from baseline to 3 weeks and 6 months after surgery using age, sex, thyroid volume, surgical procedure, and RLN/EBSLN paresis as independent variables.

The sample size was calculated to 52 patients, based on an improvement in DSQoL after surgery of 10 points in ThyPRO scores, with a standard deviation of 25 points. Cohen’s effect sizes (ES) were used, with an ES of 0.2 to 0.5 defined as small, >0.5 to 0.8 as moderate, and >0.8 as large. The study was approved by the Regional Scientific Ethical Committee for Southern Denmark (S-20130096) and registered at the Danish Data Protection Agency and at www.clinicaltrials.gov (NTC02468921).

Results

Patient Characteristics

Of 268 consecutive patients with benign nodular thyroid disease, 230 were assessed for eligibility, 62 of whom were included in the study (Figure 1). Fifty-five patients

Figure 1. Flow diagram of patient inclusion and follow-up.
completed all examinations, while 7 patients (11%) were lost to follow-up. The mean age of the patients was 52 ± 15 years (Table 1) compared with 50 ± 13 years for the patients who were eligible but not included (P = .19).

**Voice and Vocal Fold Outcome Measures**

None of the 55 patients had vocal fold changes at baseline, whereas VLS examination 3 weeks after surgery revealed uni- or bilateral paresis of the RLN in 7 patients (13%), of which 2 had paralysis of the RLN (4%). In 2 of the 7 patients and in an additional 6 patients, the ability to lengthen the vocal folds was impaired, suggesting EBSLN paresis (15%). In total, 13 patients (24%) had signs of postoperative paresis. Five patients (9%) had a localized vocal fold lesion (2 with minor edema, 2 with contact granulomas, and 1 with a small leukoplakia), of whom 1 also had RLN paresis. In total, 17 patients (31%) had VLS changes after surgery. Six months after surgery, 2 of the 7 patients with RLN paresis still had impaired vocal fold mobility (4%). However, neither of the 2 patients with initial RLN paralyses had any signs of persisting nerve injury. Two of the 6 patients with signs of EBSLN paresis still had impaired ability to lengthen the vocal folds (4%). One patient had a persistent localized vocal fold lesion (2%). Thus, a total of 5 patients (10%) had VLS changes by 6 months after surgery. Neither age, sex, thyroid volume, nor the extent of surgery had an impact on these results.

Three weeks after surgery, the patient cohort experienced an average frequency range decrease from 35 ± 5 ST to 32 ± 5 ST (P < .001, ES = 0.56), and the average intensity range decreased slightly from 65 ± 8 dB to 63 ± 7 dB (P = .03, ES = 0.27), along with a moderately reduced average maximum f0 of 3 ± 6 ST (P < .001, ES = 0.66) and an average reduced maximum SPL of 3 ± 6 dB (P < .001, ES = 0.57; Table 2). These findings reflect a reduced ability to shout/achieve loudness or to reach the highest pitch in the weeks after surgery. These parameters returned to baseline levels 6 months after surgery. No significant changes were observed in the MDVP parameters jitter, shimmer, NHR, VTI, and SPI or MPT nor in GRB scales after surgery when analyzing the entire cohort.

Postoperative paresis was associated with a significant 4.3 ± 7.5 ST decrease in the maximum f0 (P < .001), a 4.7 ± 7.9 ST decrease in the frequency range (P < .001), a 5.3 ± 8.2 dB reduction in maximum intensity, and a 4.3 ± 10.5 dB reduction in intensity range (P = .002). Paresis was also associated with a 4.5 ± 11.2 second reduction in MPT (P = .001), a 0.89 ± 1.41 increase in jitter (P < .001), a 0.01 ± 0.04 increase in NHR (P = .02), and an increase of 0.40 ± 1.19 in grade (P = .01), 0.42 ± 1.41 in roughness (P = .02), and 0.36 ± 1.11 in breathiness (P = .01; Table 3). Patients with no sign of RLN/EBSLN paresis experienced only minor changes, with difficulties reaching the highest pitch 3 weeks after surgery (Table 4).

**The Impact of Thyroidectomy on VHI**

At baseline, the median VHI score was 5 points (range, 0-70 points). Despite a statistically significant average increase of 7 ± 22 points (P = .003, ES = 0.53), it remained low 3 weeks after surgery, with a median score of 7 points (range, 0-87 points; P = .003, ES = 0.53; Table 2). The wide ranges indicate a very variable perceived voice handicap for the entire patient cohort. Thus, 16 patients (29%) had a ≥10-point increase (more complaints) in VHI score, while 7 patients (13%) experienced a ≥10-point decrease (fewer complaints) in VHI scores, despite normal voice function at baseline. Six months after surgery, the median VHI score of 9 points (range, 4-80 points; P = .30, ES = 0.18) was comparable with the baseline level. Six months after surgery, 8 of the patients (15%) had a VHI score that was ≥10 points higher than baseline values, while 11 patients (20%) had a ≥10-point lower VHI score compared with baseline. Postoperative paresis corresponded to a 12 ± 28 point increase in VHI compared with patients with no paresis (P = .002). There was a negative correlation between VHI and the maximum f0 and maximum SPL, corresponding to increased subjective voice handicap with a decreased ability to shout or to raise the voice pitch. Each reduction of 1 unit of ST in maximum f0 or 1 dB in maximum SPL led to an increased VHI score of 0.76 points (P = .02) and 1.36 points (P < .001), respectively.

**The Impact of Thyroidectomy on QoL**

The mean ThyPRO scale scores are shown in Table 5. The data at baseline and 6 months postoperatively have previously been published as part of a larger cohort investigating DSQoL in comparison with the general population.9 At 3 weeks after surgery, the Impaired Daily Life score had increased (ie, deterioration; P < .001, ES = 0.81), while most other scores showed small improvements (ES < 0.5).
Six months after surgery, symptoms were reduced significantly for all scales, except Impaired Social Life, Impaired Daily Life, and Impaired Sex Life scales. Neither the extent of surgery (hemithyroidectomy vs total thyroidectomy) nor size of the thyroid correlated significantly with the scores of the ThyPRO scales.
Postoperative RLN/EBSLN paresis was significantly associated with the Goiter Symptom scale, with an 11.0-point poorer improvement in goiter symptoms \((P = .02)\) as compared with patients with no paresis, while no association with any other ThyPRO scale was found.

Table 4. Voice Scores (Mean ± SD and Median [Range]) for the 42 Patients without Signs of RLN/EBSLN Paresis 3 Weeks after Thyroidectomy.a

<table>
<thead>
<tr>
<th>Voice Parameter</th>
<th>Baseline</th>
<th>3 wk after Surgery</th>
<th>6 mo after Surgery</th>
<th>(P) Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency range, ST</td>
<td>35 ± 5</td>
<td>34 ± 5</td>
<td>34 ± 5</td>
<td>.01</td>
</tr>
<tr>
<td>Intensity range, dB</td>
<td>65 ± 7</td>
<td>64 ± 6</td>
<td>66 ± 7</td>
<td>.54</td>
</tr>
<tr>
<td>Maximum (f_0), ST</td>
<td>69 ± 5</td>
<td>67 ± 5</td>
<td>68 ± 5</td>
<td>.&lt;.001</td>
</tr>
<tr>
<td>Maximum SPL, dB</td>
<td>114 ± 5</td>
<td>113 ± 5</td>
<td>114 ± 5</td>
<td>.01</td>
</tr>
<tr>
<td>Minimum (f_0), ST</td>
<td>34 ± 5</td>
<td>33 ± 5</td>
<td>33 ± 5</td>
<td>.02</td>
</tr>
<tr>
<td>Minimum SPL, dB</td>
<td>50 ± 5</td>
<td>49 ± 4</td>
<td>48 ± 5</td>
<td>.20</td>
</tr>
<tr>
<td>Jitter</td>
<td>0.77 (0.23-2.94)</td>
<td>0.64 (0.28-3.21)</td>
<td>0.77 (0.29-2.63)</td>
<td>.50</td>
</tr>
<tr>
<td>Shimmer</td>
<td>2.98 (1.35-7.91)</td>
<td>3.36 (1.36-6.51)</td>
<td>3.14 (1.60-5.85)</td>
<td>.08</td>
</tr>
<tr>
<td>NHR</td>
<td>0.12 (0.09-0.22)</td>
<td>0.13 (0.10-0.17)</td>
<td>0.12 (0.07-0.17)</td>
<td>.05</td>
</tr>
<tr>
<td>VTI</td>
<td>0.04 (0.02-0.10)</td>
<td>0.04 (0.03-0.08)</td>
<td>0.04 (0.03-0.09)</td>
<td>.09</td>
</tr>
<tr>
<td>SPI</td>
<td>10.06 (2.43-43.54)</td>
<td>8.43 (2.89-19.64)</td>
<td>9.16 (2.40-29.96)</td>
<td>.18</td>
</tr>
</tbody>
</table>

Table 5. Mean ± SD ThyPRO Scale Scores for Patients with Benign Nodular Thyroid Disease (n = 55).a

<table>
<thead>
<tr>
<th>ThyPRO Domain</th>
<th>Baseline</th>
<th>3 wk after Surgery</th>
<th>(P) Value</th>
<th>Effect Size</th>
<th>6 mo after Surgery</th>
<th>(P) Value</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goiter Symptoms</td>
<td>43 ± 23</td>
<td>34 ± 23</td>
<td>&lt;.001</td>
<td>0.39</td>
<td>9 ± 9</td>
<td>&lt;.001</td>
<td>1.47</td>
</tr>
<tr>
<td>Hyperthyroid Symptoms</td>
<td>23 ± 18</td>
<td>16 ± 18</td>
<td>.003</td>
<td>0.39</td>
<td>14 ± 14</td>
<td>&lt;.001</td>
<td>0.49</td>
</tr>
<tr>
<td>Hypothyroid Symptoms</td>
<td>21 ± 19</td>
<td>16 ± 20</td>
<td>.09</td>
<td>0.23</td>
<td>15 ± 18</td>
<td>.02</td>
<td>0.29</td>
</tr>
<tr>
<td>Tiredness</td>
<td>51 ± 27</td>
<td>47 ± 29</td>
<td>.42</td>
<td>0.17</td>
<td>36 ± 25</td>
<td>&lt;.001</td>
<td>0.56</td>
</tr>
<tr>
<td>Cognitive Complaints</td>
<td>20 ± 21</td>
<td>15 ± 23</td>
<td>.08</td>
<td>0.25</td>
<td>14 ± 21</td>
<td>.02</td>
<td>0.32</td>
</tr>
<tr>
<td>Anxiety</td>
<td>22 ± 22</td>
<td>14 ± 16</td>
<td>&lt;.001</td>
<td>0.33</td>
<td>14 ± 21</td>
<td>&lt;.001</td>
<td>0.60</td>
</tr>
<tr>
<td>Depressivity</td>
<td>26 ± 22</td>
<td>21 ± 18</td>
<td>.07</td>
<td>0.25</td>
<td>18 ± 25</td>
<td>.003</td>
<td>0.38</td>
</tr>
<tr>
<td>Emotional Susceptibility</td>
<td>30 ± 23</td>
<td>26 ± 24</td>
<td>.20</td>
<td>0.18</td>
<td>22 ± 22</td>
<td>.008</td>
<td>0.35</td>
</tr>
<tr>
<td>Impaired Social Life</td>
<td>7 ± 13</td>
<td>7 ± 13</td>
<td>.92</td>
<td>0.04</td>
<td>5 ± 12</td>
<td>.23</td>
<td>0.19</td>
</tr>
<tr>
<td>Impaired Daily Life</td>
<td>13 ± 18</td>
<td>28 ± 24</td>
<td>&lt;.001</td>
<td>0.81</td>
<td>7 ± 16</td>
<td>.06</td>
<td>0.31</td>
</tr>
<tr>
<td>Impaired Sex Life</td>
<td>15 ± 23</td>
<td>16 ± 24</td>
<td>.91</td>
<td>0.03</td>
<td>10 ± 23</td>
<td>.07</td>
<td>0.27</td>
</tr>
<tr>
<td>Cosmetic Complaints</td>
<td>20 ± 20</td>
<td>19 ± 16</td>
<td>.59</td>
<td>0.04</td>
<td>11 ± 16</td>
<td>&lt;.001</td>
<td>0.45</td>
</tr>
<tr>
<td>Overall Quality of Life</td>
<td>34 ± 28</td>
<td>26 ± 28</td>
<td>.03</td>
<td>0.31</td>
<td>12 ± 25</td>
<td>&lt;.001</td>
<td>0.80</td>
</tr>
</tbody>
</table>

*Data achieved at baseline and at 6 months after surgery have in part been published previously.\(^9\) Effect sizes of 0.2-0.5 were considered as small, values above 0.5 to 0.8 as moderate and above 0.8 as large.

*Comparison of patient scores (mean ± SD) between baseline and 3 weeks after surgery, using the linear mixed model.

*Comparison of patient scores (mean ± SD) between baseline and 6 months after surgery, using the linear mixed model.

*Data achieved at baseline and at 6 months after surgery have in part been published previously.\(^9\) Effect sizes of 0.2-0.5 were considered as small, values above 0.5 to 0.8 as moderate and above 0.8 as large.

*Comparison of patient scores (mean ± SD) between baseline and 6 months after surgery, using the linear mixed model.

*Comparison of patient scores (mean ± SD) between baseline and 6 months after surgery, using the linear mixed model.

*Data achieved at baseline and at 6 months after surgery have in part been published previously.\(^9\) Effect sizes of 0.2-0.5 were considered as small, values above 0.5 to 0.8 as moderate and above 0.8 as large.

*Comparison of patient scores (mean ± SD) between baseline and 6 months after surgery, using the linear mixed model.

*Comparison of patient scores (mean ± SD) between baseline and 6 months after surgery, using the linear mixed model.

*Data achieved at baseline and at 6 months after surgery have in part been published previously.\(^9\) Effect sizes of 0.2-0.5 were considered as small, values above 0.5 to 0.8 as moderate and above 0.8 as large.

*Comparison of patient scores (mean ± SD) between baseline and 6 months after surgery, using the linear mixed model.

*Comparison of patient scores (mean ± SD) between baseline and 6 months after surgery, using the linear mixed model.

*Data achieved at baseline and at 6 months after surgery have in part been published previously.\(^9\) Effect sizes of 0.2-0.5 were considered as small, values above 0.5 to 0.8 as moderate and above 0.8 as large.

*Comparison of patient scores (mean ± SD) between baseline and 6 months after surgery, using the linear mixed model.

*Comparison of patient scores (mean ± SD) between baseline and 6 months after surgery, using the linear mixed model.

*Comparison of patient scores (mean ± SD) between baseline and 6 months after surgery, using the linear mixed model.

*Data achieved at baseline and at 6 months after surgery have in part been published previously.\(^9\) Effect sizes of 0.2-0.5 were considered as small, values above 0.5 to 0.8 as moderate and above 0.8 as large.

*Comparison of patient scores (mean ± SD) between baseline and 6 months after surgery, using the linear mixed model.

*Comparison of patient scores (mean ± SD) between baseline and 6 months after surgery, using the linear mixed model.

*Comparison of patient scores (mean ± SD) between baseline and 6 months after surgery, using the linear mixed model.

*Data achieved at baseline and at 6 months after surgery have in part been published previously.\(^9\) Effect sizes of 0.2-0.5 were considered as small, values above 0.5 to 0.8 as moderate and above 0.8 as large.

*Comparison of patient scores (mean ± SD) between baseline and 6 months after surgery, using the linear mixed model.

*Comparison of patient scores (mean ± SD) between baseline and 6 months after surgery, using the linear mixed model.

*Comparison of patient scores (mean ± SD) between baseline and 6 months after surgery, using the linear mixed model.

*Data achieved at baseline and at 6 months after surgery have in part been published previously.\(^9\) Effect sizes of 0.2-0.5 were considered as small, values above 0.5 to 0.8 as moderate and above 0.8 as large.

*Comparison of patient scores (mean ± SD) between baseline and 6 months after surgery, using the linear mixed model.

*Comparison of patient scores (mean ± SD) between baseline and 6 months after surgery, using the linear mixed model.

*Comparison of patient scores (mean ± SD) between baseline and 6 months after surgery, using the linear mixed model.

*Data achieved at baseline and at 6 months after surgery have in part been published previously.\(^9\) Effect sizes of 0.2-0.5 were considered as small, values above 0.5 to 0.8 as moderate and above 0.8 as large.

*Comparison of patient scores (mean ± SD) between baseline and 6 months after surgery, using the linear mixed model.

*Comparison of patient scores (mean ± SD) between baseline and 6 months after surgery, using the linear mixed model.

*Comparison of patient scores (mean ± SD) between baseline and 6 months after surgery, using the linear mixed model.

*Data achieved at baseline and at 6 months after surgery have in part been published previously.\(^9\) Effect sizes of 0.2-0.5 were considered as small, values above 0.5 to 0.8 as moderate and above 0.8 as large.

*Comparison of patient scores (mean ± SD) between baseline and 6 months after surgery, using the linear mixed model.

Discussion

Using multiple assessment points, well-validated questionnaires, and blinded VLS examinations, this study is the first to demonstrate how signs of postoperative RLN/EBSLN
Paresis relate to the degree of symptom relief in patients with benign thyroid nodular disease. VLS is superior to indirect laryngoscopy or white-light fiberlaryngoscopy, which often underestimates the frequency of vocal fold changes after thyroidectomy. This may explain the high percentage (31%) of patients with vocal fold changes at 3 weeks after surgery in our study. Five of our patients had a minor edema, granuloma, or leukoplakia on the vocal folds, which might be related to a traumatic intubation procedure, use of large-diameter NIM tubes, or postoperative voice abuse, whereas the remaining VLS changes (signs of RLN or EBSLN paresis) are more likely associated with the surgical procedure itself.

The multifactorial voice examination showed significantly reduced mean values of the maximum $f_0$ and maximum SPL, probably because of the high number of patients with signs of RLN/EBSLN paresis who experienced a more pronounced deterioration in voice parameters than the remaining patients. Neither age, sex, goiter volume, nor surgical procedure was statistically correlated with RLN/EBSLN paresis, but this may be explained by the lack of power in our study.

A few previous studies have also examined postoperative vocal fold changes after thyroidectomy. Vicente et al identified VLS abnormalities in 21% of patients 2 weeks after total thyroidectomy but only in 6% of patients after a hemithyroidectomy. At 6 months after surgery, these figures were significantly reduced to 7% having VLS changes after total thyroidectomy, while hemithyroidectomized patients showed no VLS abnormalities. Those data lend credence to the theory that post-thyroidectomy RLN/EBSLN paresis is underestimated. In another study, Ryu et al found a decrease in maximum $f_0$ of $>10\%$ in 57% of patients 1 month after surgery, with no improvement 12 months postoperatively. An increased VHI of $\geq 10\%$ was found in 63% of patients 1 month postoperatively and in 84% 12 months after surgery. In contrast, Maeda et al found no change in subjective voice handicap using the VHI-10 questionnaire in the first month following surgery. The above-mentioned studies also included patients with thyroid malignancy, who may have a higher threshold for accepting any discomfort associated with surgery. This may explain the more pronounced subjective voice handicap reported by the patients in our study.

In line with a previously published study from our institution, we found that patients with thyroid nodularity experienced substantial symptom relief and improved QoL after thyroidectomy. Importantly, RLN/EBSLN paresis 3 weeks after surgery was associated with less improvement in the Goiter Symptoms score but did not correlate with the other ThyPRO scale scores. This association may be explained by the fact that the Goiter Symptoms score includes the item, “Have you experienced hoarseness in the last weeks?” Our data therefore support the general safety of thyroid surgery, as postoperative RLN/EBSLN paresis had only little negative impact on QoL. Importantly, RLN/EBSLN paresis seems to delay or reduce relief of the goiter symptoms. Kuhn et al retrospectively investigated the impact of post-thyroidectomy voice disorders on QoL in thyroid cancer survivors using a nonvalidated 36-item questionnaire. Fifty-one percent of patients reported post-thyroidectomy voice disorders, but the response rate was only 37%. A quarter of these patients reported a detrimental impact of the voice disorders on their QoL. These results are not directly comparable with ours because of the differences in study design. However, it is evident from the study by Kuhn et al. and supported by the present results, that postoperative voice changes affect QoL. Thus, it is crucial to inform patients about how relief of goiter symptoms may depend on the occurrence of RLN/EBSLN paresis. With nonsurgical goiter treatment, such as radiiodine or ultrasound-guided interventions, side effects are more likely avoided. How such nonsurgical interventions affect voice-related QoL is, however, at present unknown.

The strengths of our study include a low number of patients lost to follow-up, the use of validated questionnaires, and blinded VLS and GRB analyses. A few limitations need to be addressed. Although our center receives patients with benign nodular goiter from a geographically well-defined area, a degree of selection bias is inevitable, as illustrated by the inclusion of only 62 of an initial 268 patients (23%) with benign nodular thyroid disease. In addition, our study might be underpowered for analyzing the impact of age, sex, surgical procedure, and thyroid volume on RLN/EBSLN paresis and therefore cannot make conclusions about subgroups of patients. In addition, electromyography examination was not used to confirm RLN/EBSLN paresis, implying that the reported incidence of paresis might be overestimated. Moreover, we excluded patients with thyroid malignancy as well as those with overt hyperthyroidism, in order to exclude the potential effects of fear of malignancy and that of increased metabolism on the perception of symptoms and QoL. Consequently, any conclusion about the impact of RLN/EBSLN paresis needs to be restricted to the thyroid phenotype investigated.

Conclusion

Signs of RLN/EBSLN paresis after thyroidectomy were associated with less pronounced improvement in goiter symptoms in patients with thyroid nodular disease. However, thyroidectomy was associated with reduced disease-related symptoms and improved DSoQoL by 6 months after surgery. This temporary side effect should be taken into account in the dialogue with the patient referred for benign thyroid nodular disease therapy. A comparison of surgical and nonsurgical treatment modalities is awaited to further qualify recommendations for choice of therapy for future patients.

Author Contributions

Jesper Roed Sorensen, handled all aspects of the study including study design, patient inclusion, data collection, analyses and interpretation of data, writing and approval of the manuscript. Trine Printz, assistance in the design and collection of voice data,
analysis and interpretation of the voice data, revision and approval of the final version of the manuscript; Jenny Iwarsson, assistance in the design, analysis, and interpretation of the GRBAS data, revision and approval of the final manuscript; Agot Møller Grøntved, substantial part of the design of the study regarding voice parameters, interpretation of voice data, writing, revision, and approval of the final manuscript; Helle Dossing, design of the study as a whole, collection and analysis of surgical data, revision and approval of the manuscript; Laszlo Hegedüs, substantial contribution to the study design and interpretation of data, revision and approval of the manuscript; Steen Joop Bonnema, design of the study, recruitment of patients, collection and interpretation of biochemical data, revision and approval of manuscript; Christian Godballe, design of the study, collection of surgical data, interpretation of data, revision and approval of manuscript; Camilla Slot Mehlum, design of work, acquisition of voice VLS data, analysis and interpretation of data, writing, revision, and approval of the manuscript.

Disclosures

Competing interests: None.

Sponsorships: University of Southern Denmark, which primarily provided facilities for hosting the research data and help with logistical assistance. The sponsor had no part in designing, collecting, analyzing, or interpreting data.

Funding source: The Region of Southern Denmark (governmental health provider), University of Southern Denmark, and A.P. Møller Foundation. Salary for included personnel and thereby help with collection of data. No additional roles. Salary for the corresponding author. No other roles. Covering of patient reimbursements fees. No additional influence on the study.

References


