Voice and swallowing symptoms after thyroid surgery assessed using the ThyVoice online platform

Radu Mihai on behalf of ThyVoice working group*

Department of Endocrine Surgery, Churchill Cancer Centre, Oxford University Hospitals NHS Foundation Trust, Oxford, UK
Correspondence should be addressed to R Mihai: radumihai@doctors.org.uk

*ThyVoice contributors: patients recruited in this study were operated under the care of the following Consultant thyroid surgeons (in alphabetical order): Miss Julie Dunn, Mr Ashley Hey, Mr Shad Khan, Mr Radu Mihai, Mr Iain Nixon, Mr David Scott-Coombes, Mr Michael Stechman, Miss Isabel Quiroga.

Abstract

Objective: The aim of this multicentre study was to investigate the progression of patient-reported outcomes after thyroid surgery, with emphasis on voice and swallowing difficulties.

Methods: An online platform was used to collect replies to standardised questionnaires (voice handicap index, VHI; voice-related quality of life, VrQoL; EAT-10) preoperatively and at 2–6 weeks and 3–6–12 months after surgery.

Results: A total of 236 patients were recruited from five centres that contributed with median of 11 cases (range 2–186 cases). Average symptoms scores showed voice changes lasting up to 3 months: VHI increased from 41 ± 15 (preop) to 48 ± 21 (6 weeks) and returned to 41 ± 15 at 6 months. Similarly, VrQoL increased from 12 ± 4 to 15 ± 6 and returned to 12 ± 4 (6 months). Severe voice changes (VHI > 60) were reported in 12% of patients preop, 22% at 2 weeks, 18% at 6 weeks, 13% at 3 months and 7% at 12 months. Only five patients with normal preoperative voice had persistent severe voice changes after 6–12 months. Those with severe voice changes at 2 weeks (median VHI 70.5, IQR 65–81) had significant improvement by 6 months (median VHI 54, IQR 39–65) (P < 0.001).

Swallowing assessment showed a median preop score of 0 (IQR 0–3) increasing to a median of 2 (IQR 0–8) at 2 weeks and normal values afterwards.

Conclusion: The ThyVoice online platform allows the assessment of patient-reported outcome measures in thyroid surgery. Voice morbidity appears to be more frequent than commonly reported, and this risk should the quoted during informed consent. Swallowing difficulties are mild but significant in the first 2 weeks.

Introduction

Thyroid surgery is the most common endocrine operation, with over 10,000 procedures being performed annually in the UK. A large proportion of these operations (over 4000 cases per year) are captured by the national audit maintained by members of the British Association of Thyroid and Endocrine Surgeons (BAETS) (1). Voice changes represent the most common postoperative complication after thyroid surgery and can have an
impact on social and professional life and therefore are the cause in the majority of cases of litigation after thyroidectomy (2, 3).

The incidence of voice changes after thyroid surgery varies widely. Most published data are derived from centres with large workload and might not be an accurate reflection of the practice in smaller units. Anecdotal evidence suggests that some surgeons quote a risk as low as 1–2% by extrapolating the incidence of recurrent laryngeal nerve (RLN) injury, but many patients experience voice changes even when there has been no apparent injury to the RLN. In a study of 395 patients operated in a single unit, 49% had voice impairment immediately after surgery (4) The BAETS audit reported that the incidence of voice changes is 6.1% after first-time surgery and 7.7% after redo surgery, but this might still be an underestimate due to the bias created by the fact that it relies on self-reported subjective assessments made by the operating surgeons rather than on assessments made by individual patients (1). Interestingly, in a study of 86 patients, voice assessment performed by consultants and senior trainees had a positive predictive value for vocal cord paralysis of only 55%, demonstrating that subjective voice assessment by clinicians is inaccurate (5).

In contrast with the relatively low figures quoted by clinicians, an international survey of over 2300 patients with thyroid cancer organised by the Thyroid Cancer Alliance showed that 36% of patients experienced unfavourable voice outcomes (6). Very similar findings were reported based on replies from 4426 members of the Thyroid Cancer Patients Association, showing that 51% of responders had post-thyroidectomy voice dysfunction such as loss of loudness and an inability to shout or sing (7). Similarly, a study completed in our unit showed that a third of patients had mild voice problems 1 year or more after surgery (8).

The conflicting figures between outcomes reported by surgeons and patients’ groups could be due to the fact that poor results are unlikely to be published, negative outcomes are rarely made public and the surgical community as a whole is far from reaching the ideal results reported in units that set the standards for the speciality. It remains possible that more subtle voice changes may be imperceptible to clinicians who are not familiar with patients’ voice range but may be very significant for patients themselves. There is therefore a real need to explore the broader spectrum of patients’ own perception of voice morbidity after thyroid surgery in order to provide informed consent to future patients.

An additional limitation of such studies is that the method for voice assessment impacts on the incidence of abnormalities detected. Studies using computerised voice analysis have reported that the majority of patients experience alteration in maximum phonation frequency range and vocal jitter (9), mean decrease of the speaking fundamental frequency and a more monotone speech (10), but it is not known if such parameters can be perceived by patients. Listening by a professional voice therapist while reading a standard text is the best assessment of voice changes based on four voice parameters (roughness, breathiness, asthenicity and strain), but this technique is restricted by the limited access to specialised voice therapists and therefore cannot be applied to large cohort studies. In this context, two standardised validated questionnaires, voice handicap index (VHI) and voice-related quality of life (VrQoL) questionnaires have been developed for the assessment of patients’ perception of subjective voice changes (11).

In contrast with voice changes, until more recently, swallowing difficulties have not been recognized as a relatively common complication after thyroid surgery. In a study of 110 patients, the authors introduced a Swallowing Impairment Score (SIS) and found that SIS was significantly worse 1 week but not 1 month and 3 months after thyroidectomy (12).

The aim of this study was to determine the incidence of voice changes in the first year after thyroid surgery using an online platform that facilitates patient-reported outcomes using validated questionnaires for voice and swallowing symptoms.

Methods

A prospective multicentre cohort study was designed to collect patient-reported outcome measures after thyroid surgery. The study was approved by the national ethics committee review (IRAS ID 197725, REC 16/EM/0070).

Patients were considered for the study if they were adults (age range 18–90 years) diagnosed with a benign or malignant thyroid disease needing thyroid surgery (lobectomy or total thyroidectomy, with or without lymph node surgery) and with the participant being willing and able to give informed consent for participation in the study. Patients were excluded from the study if they were unwilling or unable to give informed consent, were unable to complete questionnaires in the English language and had previous thyroid surgery.
An online platform (https://thyvoice.ndorms.ox.ac.uk/) was created to collect replies to standardised questionnaires (vide infra). After recruitment by individual clinicians during preoperative clinic appointment, patients received an automatic email to allow registration with the online platform, to confirm their consent and to get access to the preoperative assessments. Once the date of the operation was confirmed by clinicians on the online database, patients received automatic messages at 2 weeks, 6 weeks, 3 months, 6 months and 12 months after surgery.

Patient-reported outcomes

1. VHI questionnaire was designed to generate three subscores for functional, physical and emotional impact of voice changes and a total score (13). A change of 18 points in a total score or 8 points on any subscale is considered statistically significant. The self-reported severity of voice changes quantified as mild, moderate and severe is reported to correlate with changes in VHI scores (Table 1).

2. VrQoL questionnaire comprises 10 items divided into physical functioning and social-emotional functioning subscales (11). Each item is scored on a 5-point interval scale that reflects the severity of the problem. The overall VrQoL score ranges from 10 to 15 (excellent), 16 to 20 (very good), 21 to 25 (good), 26–30 (fair) and 30–50 (poor).

3. EAT-10, a validated questionnaire for self-perceived symptoms of oropharyngeal dysphagia, has excellent internal consistency and reproducibility for assessment of swallowing problems (14). A summated EAT-10 total score ranges from 0 to 40, with a score ≥ 3 indicative of oropharyngeal dysphagia and a score > 15 indicative of aspiration risk (15).

Statistical analysis

Data were downloaded anonymously without access to patients’ contact details; therefore, missing replies could not be requested by contacting the patients individually.

Before analysis, data were scrutinised and incomplete replies were removed (e.g. when the sum of individual elements of a score led to a score lower than the minimum range demonstrating that some questions were not answered).

Data with normal distribution are presented as mean ± s.d. and compared using t-tests. For all other data, median (IQR) is presented and data were compared with nonparametric tests.

All analysis was done using StatPlus programme (StatPlus:mac, AnalystSoft Inc, Version v8; https://www.analystsoft.com/en/).

Results

Clinical parameters

Between August 2020 and May 2022, a total of 236 patients were recruited from five surgical centres that contributed with a median of 11 cases (range 2–186 cases) (Fig. 1). There were 164 women and 72 men of mean age 58 ± 12 years. Patients were non-smokers (65%), previous smokers (21%) or current smokers (14%). A third of patients declared to use their speaking voice for the job (e.g. vicar, teacher and salesperson). The singing voice was used either for the profession (5%) or for activities outside work (e.g. choir, church and hobby) (6%).

Patients had thyroid lobectomy (n = 133, 56%) or total thyroidectomy (n = 103, 44%) for multinodular goitre (n = 90, 38%), Graves’ disease (n = 55, 23%), thyroid nodule with suspicious THY3 cytology (n = 50, 21%), thyroid cancer (n = 37, 16%) or thyroid cyst (n = 4, 1%). A minority of patients had central compartment lymph node dissection (n = 5, 2%) or central and lateral compartment lymph node dissection (n = 21, 9%). Only 81 patients had a preoperative laryngoscopy, and a single patient was identified to have a vocal cord palsy (symptomatic patient, with locally advanced thyroid cancer). Intraoperative nerve monitoring (IONM) was used in the majority of cases (n = 170, 72%) and an abnormality was detected in three patients (1%). The main operating surgeon was a...
consultant surgeon in two-thirds of cases and a fellow or an advanced surgical trainee in a third of cases.

Voice outcomes

Preoperatively, using a visual analogue scale (VAS) from 1 (normal voice) to 10 (severely impaired voice), only 68 of 150 (45%) responders described their preoperative voice to be completely normal and 14/150 reported severe voice changes (score 8–10) (Fig. 2A). Based on the VHI scores, 69/139 (50%) had normal voice (VHI < 35) and the others had voice symptoms quantified as mild (n = 34 (24.5%), VHI 35–44), moderate (n = 19 (13.6%), VHI 44–60) or severe (n = 17 (12.2%), VHI > 60) (Fig. 2B).

There was a weak correlation between the VAS scores and the VHI scores: patients with low VAS scores (1–2) had lower VHI scores compared with those with high VAS scores (8–10) (35.4 ± 8.2 vs 52.8 ± 18.9, P < 0.01 t-test) (R=0.446). In contrast, there was a very good correlation between VHI and VrQoL score (R=0.819) (Fig. 3).

For the entire cohort, compared with the preoperative VHI scores (40.9 ± 14.5), there was a significant increase in VHI scores at 2 weeks (49.4 ± 19.9, P < 0.01) and 6 weeks (48.4 ± 20.7, P < 0.01), but this difference disappeared at 3 months (43.2 ± 16.2, P=NS), 6 months (41.2 ± 15.4, P=NS) and 12 months (41.4 ± 14.3, P=NS) (comparison with preoperative values analysed using t-test for pairs) (Fig. 4A). Same pattern was observed for VrQoL scores (Fig. 4B).

There was no significant difference in VHI scores between patients who had lobectomy or total thyroidectomy at preoperative assessment (41.9 ± 15.2 vs 39.9 ± 14.0) nor at 2 weeks (49.1 ± 18.6 vs 50.0 ± 23.8) or 6 weeks (48.2 ± 19.6 vs 48.9 ± 22.5). The same pattern
was observed for VrQoL scores (12.9 ± 4.8 vs 11.9 ± 3.3 preoperatively, 14.4 ± 5.9 vs 15.5 ± 6.9 at 2 weeks and 13.8 ± 5.4 vs 14.7 ± 6.5 at 6 weeks).

Of the 69 patients with normal preoperative voice (VHI < 35), 19 patients reported worse scores at 2 weeks (45.2 ± 20.9) and 6 weeks (44.5 ± 20.3) but with a return towards normal scores at 3 months (38.7 ± 9.5) and 6 months (37.6 ± 11.7). Only one patient had severe voice changes at 6 months (VHI 75) and 12 months (VHI 60). Interestingly, one patient with very severe voice symptoms at 2–6 weeks (VHI 107–112) had normal voice at 6 months (VHI 30).

Of the 17 patients with preoperative voice changes quantified as severe (VHI > 60), all but one patient improved gradually, with the scores decreasing from 74.6 ± 8.1 preoperatively to 66.2 ± 21.9 at 2 weeks, 65.7 ± 26.9 at 6 weeks (P = NS) and 55.3 ± 22.9 at 6 months (P < 0.01).

Severe voice changes (VHI > 60) were reported in 12% of patients preop, 23% at 2 weeks, 18% at 6 weeks, 13% at 3 months, 11% at 6 months and 7% at 12 months.

Swallowing symptoms

Preoperatively, 83/148 (56%) patients had no swallowing symptoms (EAT-10 score = 0) and 7 (8%) of these patients reported significant symptoms at 6–12 months after the operation (EAT-10 scores 3–11). Severe swallowing problems were reported preoperatively by 47/149 (32%) patients with a median EAT-10 score of 6 (IQR 4–11), and all but 3 patients improved by 6–12 months postoperatively.

For the entire cohort, there was a significant increase in median EAT-10 scores from 0 (0–3) preoperatively to 2 (0–8) at 2 weeks (P < 0.05, ANOVA), but at all other time points, the scores were similar: 1 (0–5) at 6 weeks, 1 (0–3) at 3 months, 0 (0–3) at 6 months and 1 (0–4) at 12 months (all comparison with preoperative scores analysed with ANOVA). The extent of surgery did not impact on the swallowing symptoms, scores at 2 weeks postop being similar in those who had lobectomy or total thyroidectomy (median 2 (IQR 0–8) vs median 2 (IQR 0–9), respectively).

Impact of intraoperative nerve monitoring

Subgroup analysis was made based on whether the IONM was used (n = 87) or not (n = 53). Preoperative scores were similar for VHI questionnaire (38.7 ± 13.3 vs 42.4 ± 15.2, P = NS) and EAT-10 questionnaire (median 0 (0–2.5) vs 0 (0–2), P = NS). There was no difference in VHI scores at 2 weeks (49.4 ± 21.5 vs 46.9 ± 17.3) and 6 weeks (47.7 ± 20.3 vs 47.8 ± 19.3).

Three patients had a loss of IONM signal, one intermittent (patient with Graves’ disease) and two definitive (sectioning of RLN in patients with retrosternal goitre and locally advanced thyroid cancer). Only two of these patients returned their postoperative questionnaires (Fig. 5). The patient with intermittent LOS had a full return to normal for the voice and swallowing scores. The patient with RLN transection showed persistent (though improving) voice symptoms and a normal swallowing score, suggesting a functional benefit of the nerve repair performed intraoperatively.

Discussion

This paper describes patient-reported voice and swallowing outcomes within 12 months after surgery in a cohort of adult patients operated in five surgical units in the United Kingdom. Similar outcome data have been published previously, but this study brought a new approach through the use of an online platform designed to allow patients to submit responses to standardised questionnaires for voice and swallowing symptoms at fixed time intervals.

The cohort is representative for the practice of thyroid surgery in the United Kingdom in terms of gender distribution, age, indication for surgery and extent of surgery when compared with figures published in the most recent report from the United Kingdom Registry of Endocrine and Thyroid Surgery (1). Despite the low number recruited in some of the centres, all surgeons involved in
the study were considered high volume, having an annual workload above the threshold recommended by the European Society of Endocrine Surgeons of over 50 thyroid operations per year (16). All patients had open surgery as robotic or transoral thyroidectomies are not performed by any of the participating surgeons. Better voice outcomes after robotic surgery were reported in some (17, 18) but not all publications (19, 20), and therefore the findings of this analysis might change if these new techniques will be adopted widely in future years.

Based on the incidence of benign (e.g. Graves’ disease) and malignant thyroid pathology in the population, there was a wide age spectrum. Because the study required participants to be confident with using the online technology, it is likely that older patients were reluctant to be recruited in the study, and this might have created an unavoidable bias. This issue is important because some studies reported that age > 50 years is independently associated with the development of voice or swallowing changes after thyroidectomy, despite intact RLN (21), and therefore, it might be that the outcomes would be skewed if older patients were systematically excluded from the study.

Either by self-assessment using a VAS or by answering the VHI questionnaires, only half of the patients needing thyroid surgery were found to have a normal voice before their operation. This high incidence of preoperative voice symptoms was unexpected, and it was possibly related to the fact that a significant proportion had retrosternal goitres or thyroid cancer. Unfortunately, the study did not record the clinicians’ impression about the preoperative voice. It is likely that the vast majority of these patients had an apparently normal voice yet, given the opportunity, they would declare symptoms that skew their preoperative VHI score towards the symptomatic range. Furthermore, the lack of correlation between the VAS scores and VHI scores suggests that patients’ own overall impression of voice performance might not pick up symptoms that become relevant once scored as part of a standardised questionnaire like VHI or VrQoL. These findings highlight the importance of a formal preoperative voice assessment as a comparative ‘individual benchmark’ for postoperative voice outcomes.

The responses to the preoperative questionnaires were not known to the surgeon on the day of the operation; therefore, patients’ perception of voice did not influence the decision to perform preoperative laryngoscopy. As many patients included in the study were operated during the coronavirus disease pandemic, there was an active decision to minimise the use of laryngoscopy both preoperatively and postoperatively. In this context, the analysis of patient-reported voice outcomes was not correlated with findings on laryngoscopy. In addition, the results of postoperative laryngoscopy performed by individual surgeons were not collected because the study was designed to be patient driven and to minimise the input of the recruiting surgeons. Once details of the operation were provided by the surgeon, there was no requirement that the surgeon will provide follow-up information and the running of the study relied on patients’ involvement. In this context, it was assumed that patients with normal IONM signals at the end of the procedure had a normal postoperative function of the RLN, an assumption based on the vast volume of published data demonstrating that IONM has a good positive predictive value for postoperative nerve function (22, 23, 24).

In patients with normal preoperative voice, a quarter had a degree of voice changes lasting up to 6 weeks but...
settling within 3–6 months. In patients presenting with severe preoperative voice changes (as suggested by VHI score > 60), the likelihood of significant improvement postoperatively was very high, with only 1 in 17 such patients in this study being found to have persistent severe voice problems at 6 months after the operation. This information should be used to reassure patients who are concerned of perioperative voice changes.

In contrast with voice changes, until more recently, swallowing difficulties have not been commonly recognized as a potential complication after thyroid surgery. In a study of 110 patients, the authors introduced a swallowing symptoms score SIS and found that SIS was significantly worse 1 week after thyroidectomy but not 1 month and 3 months after thyroidectomy (12). Same fast recovery of swallowing symptoms within 1 month after thyroid surgery was demonstrated in a prospective study showing that the length of laryngotracheal elevation measured on ultrasound was related to the symptom score of swallowing after thyroid surgery and therefore suggesting that the presence of adhesion between the laryngotracheal unit and the superficial soft tissue is the probable cause of such symptoms (25). This reported fast improvement in swallowing symptoms mirrors the findings of this study. Based on the scores on the EAT-10 questionnaire there, swallowing appeared to be affected only at the 2-week assessment and it showed no significant changes at later time points.

There are limitations of the current study that could not have been avoided.

First, the contribution of each centre varied widely and, in the absence of significant multicentre input, it remains difficult to use this data as a benchmark of the current standard of thyroid surgery. At the launch of the study, 40 centres showed an interest in contributing and 32 centres started the process of registering as a contributor, of who 7 centres were approved and 5 centres started recruiting patients. As the study was not enrolled on the NIHR portfolio because it was not founded through a nationally competed grant, the motivation of individual contributors waned. In this context, the initial aim of the ThyVoice study of recruiting 1000 patients was revised, and the study was concluded within 2 years.

The second limitation was the variability of engagement of individual patients with the study, with majority of them responding only to some of the planned assessments. In particular, patients ‘abandoned’ the study once they had no symptoms at the previous assessment. As participation was entirely voluntary, there was no mechanism to approach patients who failed to respond to the automatic reminders at each time point of the study. This led to a relatively high number of missing datapoints that could not be corrected.

The strength of this study was the ability to collect long-term outcome data outside the routine time scale for clinical follow-up appointments. The online access to the questionnaires and the automatic reminders to complete the questionnaires at designed time points was expected to provide an accurate evaluation that would be impossible to secure through face-to-face interviews during clinic appointments. Patient engagement with the study was crucial for data collection as there was no mechanism to request replies from those who failed to respond to the automatic reminders. Convenient and efficient as the online platform might be, it has the drawback of not being accessible by those who are less comfortable with internet usage. As this limitation becomes less predominant within the society, this established ThyVoice platform could be used in future prospective studies that would include patient-reported outcome data.

In summary, the ThyVoice online platform is an effective mechanism for collecting patient-reported outcomes after thyroid surgery, and its use could be extended to future studies on this topic. Voice symptoms are very common in the first 2–6 weeks after thyroidectomy but are very unlikely to persist over 6–12 months. Immediately after the operation, swallowing symptoms are more frequent than anticipated, but they settle within 6 weeks or less. The data showed no evidence that the use of IONM can mitigate such symptoms. Overall, this study confirms previous publications on the topic and adds a timeline for the likelihood of voice and swallowing symptoms after thyroid surgery that can be used when obtaining the informed consent for thyroid surgery.

Declaration of interest
There is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding
A research grant from the British Association of Endocrine and Thyroid Surgeons (BAETS) allowed the development of the IT infrastructure. Administrative support was funded through a grant from the Research Capability Funding – Oxford University Hospital NHS Foundation Trust.

Ethics approval
The study was approved by the national ethics committee review (IRAS ID 197725, REC 16/EM/0070, East Midlands – Nottingham 1 Research Ethics Committee).
Acknowledgements

The author is grateful to Joe Barrett and David Gray for the IT expertise to establish and maintain the online ThyVoice platform and for their administrative help. The following surgical trainees and Fellows contributed to data collection (in alphabetical order): Mr Vishnusai Chauhan, Miss Machteld de Jong, Mr Neil Patel, Miss Brooke Puttergill, Miss Elina Shaari, Miss Vi Sia and Miss May Thwin. Data were presented at the BAETS2022 meeting (Bournemouth, October 2022).

References