The impact of laryngeal biopsy on voice outcomes: a pilot study

Liza Bergström1,2,*, Elizabeth C Ward2,3 and Caterina Finizia1

1Department of Otorhinolaryngology, The Sahlgrenska Academy, University of Gothenburg, Sweden
2School of Health & Rehabilitation Sciences, The University of Queensland, Australia
3Center for Functioning & Health Research (CFAHR), Queensland Health, Australia

Abstract

Objective: Several studies have suggested that type of laryngeal biopsy may impact voice and vocal function, which may ultimately influence functional outcomes post cancer treatment. This study directly examined the impact of the biopsy procedure on voice outcomes for early glottic (Tis-T2) cancer patients.

Design: Prospective pilot cohort study.

Method: Fifteen patients diagnosed with early glottic cancer underwent voice recordings within one week pre-biopsy and between one-six weeks post-biopsy, via Sahlgrenska University Hospital, Sweden. A control group (n=15), matched for gender and comparable for age and smoking status, was assessed once only. Multidimensional voice analyses were conducted with both groups including: (1) Grade, Roughness, Breathiness, Asthenia, Strain (GRBAS) perceptual rating, (2) acoustic measures of harmonics to noise ratio, jitter, shimmer, mean spoken fundamental frequency and (3) Maximum Phonation Time (MPT).

Results: In comparison to the control group, most perceptual and acoustic parameters were significantly (P=0.05) more impaired both pre and post biopsy. No significant difference (P>0.05) in the patient cohort was observed between pre-post biopsy voice parameters. At an individual level, half of the patients showed a perceived change in voice post biopsy, of which four demonstrated improvement and three a deterioration in vocal function.

Implications: Although group level analysis did not show a significant biopsy-voice impact, individual data suggests that multiple, punch biopsies may negatively impact functional outcomes. The biopsy procedure is undoubtedly a necessary first step in laryngeal cancer management; however, given the emerging discussion regarding the type of biopsy technique, the nature and the extent of tumour/tissue removal on functional outcomes, this is an area for further research.

Introduction

Larynx cancer is the second most common malignancy of the respiratory tract and of all head and neck cancers [1,2], with glottic carcinoma the most common type [3]. To date the majority of research on voice outcomes following early glottic cancer (Tis-T2) has explored the impact of either radiotherapy or transoral laser microsurgery on voice quality and functional outcomes [4-6]. Findings from this body of research suggests that both methods of organ-preservation treatment impact vocal quality and function to a similar extent, with the majority of patients expected to present with persistent, predominantly mild-moderate changes to voice quality [7-9]. Whilst numerous mechanisms relating to either the radiotherapy or surgical procedures have been proposed to be the cause of such vocal change, few studies have explored the potential impact of the earliest procedure on patient outcomes, that of the biopsy [10-13].

The biopsy, an important key step in the initial diagnostic process for laryngeal cancer, can occur using a variety of techniques and is dependent upon lesion characteristics (superficial versus invasive versus exophytic), location of the lesion and individual surgeon skill and preferences [1,14,15]. For glottic tumours, the endoscopic excisional biopsy is favourably reported since it has the dual purpose in that it is both diagnostic and optimises therapeutic outcomes in a single procedure [14,16,17]. An excisional biopsy removes the entire pathological tissue whilst endeavouring to preserve structures of the vocal cords and is argued to be superior to the punch biopsy which may harm the vocal cords [14]. Vocal cord stripping (removal of the top most layers of tissue along the vocal cords) may also occur as a biopsy procedure [12,13,18].

Numerous authors lay claim that the biopsy process may influence voice and functional outcomes [11-13,18], however the possible biopsy-voice impact has not been systematically investigated in a single study to date. All studies, discussing the impact of the biopsy, have been retrospective and were designed to explore post cancer treatment voice outcomes, with the potential for possible impact from the biopsy procedure only discussed as a post hoc, secondary measure. In addition, no study has compared voice function at the specific time-points of pre and post biopsy.

However despite these limitations, authors argue that the greater the excisional tumour/tissue removal the greater the grade of dysphonia [17-19]. Others propose that the process of stripping of the vocal cords versus a simple biopsy procedure may result in greater deviant voice characteristics [12,18]. In contrast, some authors suggest no difference

Correspondence to: Liza Bergström, Centre for Functioning and Health Research (CFAHR), Queensland Health, Australia; Tel: +61 7 38963081; E-mail: liza.bergstrom@uq.net.au

Key words: glottic cancer, diagnosis, vocal function

Received: May 25, 2016; Accepted: May 31, 2016; Published: June 02, 2016
in long term post treatment voice outcomes for patients who had undergone stripping versus a simple biopsy procedure [11,13]. Finally, several studies now suggest that the excisional biopsy should be the procedure of preference given its dual diagnostic and therapeutic purpose [14,16,17] and the potential to preserve vocal fold function [14].

Currently there is an absence of any direct evidence that has examined the impact of the biopsy procedure on voice outcomes. If, indeed, a certain type or specific biopsy process was a potential factor in contributing to eventual post treatment voice outcomes, such information would be beneficial for improving biopsy technique for best clinical practice. Given this, and the paucity of systematic, well-designed studies, this is an area in need of further research. This pilot study aims to prospectively investigate the biopsy procedure and its impact, if any, on voice outcomes for a cohort of early glottic cancer patients (Tis-T2).

Method

Ethical considerations

This study is part of a larger clinical research trial at Sahlgrenska University Hospital, Gothenburg, Sweden, exploring voice outcomes following non-surgical laryngeal cancer management. Ethics have been approved by the Regional Research Ethics Committee, University of Gothenburg, Sweden and conducted in accordance with the Declaration of Helsinki. All participants have given their informed consent.

Study population

Patients diagnosed with laryngeal cancer in the Västra Götaland Region were referred to a weekly tumour conference at the Otorhinolaryngology Department, Sahlgrenska University Hospital. From here they were recruited into the larger trial. As part of that protocol, all consenting patients underwent voice assessments at multiple time points relative to treatment, including pre and post biopsy. For inclusion in the current investigation, only those patients who had; (1) been diagnosed with early (Tis-T2) glottic cancer; (2) completed pre-biopsy voice recordings within one week of the biopsy procedure, and; (3) completed voice recordings post-biopsy between the second and sixth week post biopsy, were included in this pilot study. Of the 32 patients with pre post biopsy data, 17 did not meet inclusion criteria, secondary to; (1) one or several biopsies occurring before the first voice recording, n=6, and (2) patient voice recordings outside of above inclusion time-frames, n=11. The final cohort within this pilot study consisted of 15 patients (13 male, two female, 53% smoking, mean age 59 years, SD 11.44, range 47-79) — see Table 1. Participants had received a range of biopsy procedures (Table 1), reflective of the current clinical practice. For two participants, the nature of the biopsy operation was unavailable. All presented perceptually as having some degree of dysphonia pre biopsy. One third of the group presented with a severe grade of dysphonia, one third presented with moderate dysphonia and one third, a degree of mild dysphonia pre biopsy. A healthy control group (n=15) was recruited for comparison which was matched for gender, comparable for age (61 years, SD 8.29, range 47-79) and proportion of smokers (47%). Statistical comparisons confirmed age (t=0.59) and proportion smoking status (χ²=0.13, P=0.72) were not significantly different between the cohorts. The absence of laryngeal pathology within the control group was confirmed by an Otolaryngologist via nasolaryngoscopy.

Voice recordings

Voice recordings, for both the participant and control group, consisted of the reading of a standard passage and the maximum sustained vowel /a/, repeated three times. A headset microphone (Sennheiser MKE 2-p) was set at a distance of 12 cm from the corner of the mouth. Recordings were made at a sampling frequency of 44.1 kHz with a Panasonic Professional Digital Audio Tape (DAT) Recorder SV-3800. Prior to analysis, all recordings were transferred from a DAT to a computer hard drive as an audio file (.wav) using the program Swell Soundfile Editor, version 4.5 (Saven Hightech).

Perceptual analyses

Perceptual ratings were conducted by two speech-language pathologists (SLP), and a third for consensus rating. Raters attended a half-day’s consensus training based on Iwarsson and Petersen (2012) [20] and anchor samples were produced and incorporated into the final rating file. The final rating file was compiled using an excerpt from each control person’s voice recording and each patient’s pre- and post-biopsy voice recording. This excerpt (.wav audio file)

<table>
<thead>
<tr>
<th>n=15</th>
<th>Gender M= Male F= female</th>
<th>Age at biopsy</th>
<th>Biopsy location and type</th>
<th>Tumour size and description</th>
<th>Dysphonia severity pre biopsy</th>
<th>Dysphonia severity post biopsy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>M</td>
<td>60</td>
<td>Bilateral punch biopsies#</td>
<td>Tis</td>
<td>1</td>
<td>2*</td>
</tr>
<tr>
<td>2</td>
<td>M</td>
<td>56</td>
<td>Left punch biopsy</td>
<td>T1a invasive tumour</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>M</td>
<td>79</td>
<td>Right punch biopsy</td>
<td>T1a verrucous tumour</td>
<td>3</td>
<td>2*</td>
</tr>
<tr>
<td>4</td>
<td>F</td>
<td>51</td>
<td>Right punch biopsy</td>
<td>T1a tumour encasing entire VF</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>M</td>
<td>56</td>
<td>Bilateral punch biopsies#</td>
<td>T1b, left VF polyp-like lesion, right leukoplakia</td>
<td>1</td>
<td>3*</td>
</tr>
<tr>
<td>6</td>
<td>M</td>
<td>64</td>
<td>Left excisional biopsy</td>
<td>T1a tumour growth along half of left VF</td>
<td>2</td>
<td>1*</td>
</tr>
<tr>
<td>7</td>
<td>M</td>
<td>71</td>
<td>Left excisional biopsy</td>
<td>T1a tumour growth superior anterior 1/3 VF</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>M</td>
<td>50</td>
<td>Right excisional biopsy</td>
<td>T1a</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>M</td>
<td>71</td>
<td>Left excisional biopsy</td>
<td>T1a</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>M</td>
<td>46</td>
<td>Left VF biopsy, not described</td>
<td>T1a tumour</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>M</td>
<td>41</td>
<td>Left punch</td>
<td>T2 tumour encasing left VF</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>12</td>
<td>M</td>
<td>63</td>
<td>Right punch biopsies#</td>
<td>T2 tumour encasing right VF</td>
<td>3</td>
<td>3*</td>
</tr>
<tr>
<td>13</td>
<td>M</td>
<td>76</td>
<td>Left excisional biopsy</td>
<td>T2 large exophytic tumour</td>
<td>2</td>
<td>1*</td>
</tr>
<tr>
<td>14</td>
<td>F</td>
<td>48</td>
<td>Left excisional biopsy</td>
<td>T2 tumour tissue encasing entire VF</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>15</td>
<td>M</td>
<td>53</td>
<td>Stripping of right VF</td>
<td>T2 right tumour encasing anterior 2/3 VF</td>
<td>3</td>
<td>2*</td>
</tr>
</tbody>
</table>

Note: VF = vocal fold, * indicates perceived change in voice quality pre post biopsy, # denotes several biopsy samples taken.

included the first two sentences of the standard passage and the second recorded prolonged vowel /a/. Fifteen percent (7/45) of these samples were randomly chosen to be reduplicated for intra-rater reliability calculations. All samples (n=52) were then randomly compiled with the anchor samples interspersed, at every 20 voice samples, into the final rating file on a USB. The raters were blinded to participant/ patient status and voice sample information. The rating protocol used the GRBAS scale [21], which consists of 5 voice qualities: Grade (G), Roughness (R), Breathiness (B), Asthenia (A), and Strain (S). Each voice quality is rated on a 4-point scale, where 0 = normal, 1 = mildly impaired, 2 = moderately impaired and 3 = severely impaired. Inter and intra-rater reliability were calculated for the two raters using percent exact agreement (PEA), percent close agreement (PCA: one-point difference) and Weighted Kappa, interpreted using Landis and Koch guidelines [22]. Inter-rater reliability revealed a PEA of 53%, and PCA of 93%. Weighted Kappa was calculated at 0.66, indicating a substantial agreement. Intra-rater reliability revealed PEA 72%, and PCA 98% and a Weighted Kappa of 0.87, indicating an almost perfect intra-rater agreement. Where ratings differed between the two clinicians, a third clinician rated the parameter and consensus rating (two of three) was used in the analysis.

Acoustic analyses

Voices were analysed using Voxalys 1.3 (Voxalys AB), a plugin programme to Praat [23]. Jitter, shimmer (perturbation measures which refer to the acoustic signal’s cycle-to-cycle variation in the fundamental frequency and amplitude, respectively) and Harmonics to Noise Ratio (HNR) values were analysed from two seconds of the middle of the second sustained vowel /a/. Mean speaking fundamental frequency (MSFF) was measured from reading of the standard passage. Aerodynamic analysis consisted of the Maximum Phonation Time (MPT) which is the longest recorded time (in seconds) for the sustained vowel /a/.

Statistics

Descriptive statistics using mean and standard deviation were calculated for perceptual, acoustic and aerodynamic voice measures. Statistical analyses were conducted using SPSS statistics software, version 21, except for the Weighted Kappa which was conducted using STATA 13. Significance was set at P<0.05. Due to small sample sizes and non-normal data distributions within the cohorts, non-parametric statistics were used for group comparisons. Differences pre to post biopsy in the patient cohort were analysed using the Wilcoxon Signed Rank Test. Comparisons between the patient and healthy control cohorts were analysed using Mann-Whitney U. In addition to group analysis, individual patterns were examined. The proportion and characteristics of individuals who demonstrated change in overall severity of dysphonia of one or more grade point on the GRBAS were noted.

Results

Perceptual voice outcomes

Table 2 reports the perceptual voice outcomes for the healthy control group, and the patient cohort at both pre and post biopsy. Relative to the control group, the patient group presented with significantly higher overall Grade of dysphonia, Roughness, Breathiness and Strain prior to biopsy. Post biopsy, the patient cohort remained significantly more impaired than controls for Grade, Breathiness and Strain. Roughness was no longer significantly different to the controls post biopsy. Visual examination of the pre-post biopsy patient group means, revealed Roughness to be the only parameter to show some degree of overall improvement pre to post biopsy, however the within-group analysis failed to support a significant change in GRBAS parameters post biopsy (Table 2).

Acoustic and aerodynamic analyses

Pre biopsy, statistically significant differences were observed between the larynx cancer patients and controls on HNR, jitter, shimmer and male MSFF (Table 3). Post biopsy, only shimmer and male MSFF continued to show a significant difference to the control group. Visual inspection of pre-post biopsy means suggest a pattern of slight improvement of HNR and slight decline in jitter post biopsy, however no acoustic parameter was found to be statistically different pre to post biopsy (Table 3). The aerodynamic measure of MPT did not differ significantly pre to post biopsy or when compared to controls.

Exploratory analyses

Although no significant pre-post biopsy changes were demonstrated at a cohort level, individually, seven patients (Table 1) showed a perceived change in voice quality pre-post biopsy. Of these seven, four individuals (patients 3, 6, 13 and 15) showed an improvement in their level of dysphonia. Examination of patient characteristics, revealed no dominant patterns regarding size of tumor nor type of biopsy procedure for these four relative to the rest of the group. The remaining three individuals (patients 1, 5 and 12) demonstrated deterioration in vocal function post biopsy. All three underwent multiple biopsies, all punch, with two of the three perceived to deteriorate to severe dysphonia post multiple biopsies.

Discussion

This is the first study to prospectively investigate the biopsy procedure and its direct impact on voice function for patients with early glottic cancer (Tis-T2). By comparing the patient data to a control cohort, the current study has confirmed that patients with Tis - T2 size tumors have significant perceptual and acoustic voice impairments pre biopsy and continue to remain with a pathological voice post biopsy. The primary reason for this dysphonia would likely be the negative impact of the tumor growth on normal vocal fold function. Post

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Healthy Controls mean (SD)</th>
<th>Patient pre-biopsy mean (SD)</th>
<th>Patient post-biopsy mean (SD)</th>
<th>Pre-post biopsy P-value</th>
<th>Control vs Patient pre-biopsy P-value</th>
<th>Control vs Patient post-biopsy P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade</td>
<td>0.53 (0.52)</td>
<td>2.00 (0.85)</td>
<td>2.00 (0.85)</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Roughness</td>
<td>0.33 (0.49)</td>
<td>1.27 (0.88)</td>
<td>0.93 (1.10)</td>
<td>0.212</td>
<td>0.003</td>
<td>0.130</td>
</tr>
<tr>
<td>Breathiness</td>
<td>0.20 (0.41)</td>
<td>1.27 (1.03)</td>
<td>1.47 (0.99)</td>
<td>0.317</td>
<td>0.001</td>
<td>0.000</td>
</tr>
<tr>
<td>Asthenia</td>
<td>0.00 (0.00)</td>
<td>0.00 (0.00)</td>
<td>0.07 (0.26)</td>
<td>0.317</td>
<td>1.000</td>
<td>0.317</td>
</tr>
<tr>
<td>Strain</td>
<td>0.20 (0.41)</td>
<td>1.53 (0.99)</td>
<td>1.53 (0.74)</td>
<td>0.957</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Note: Significance = P<0.05, indicated by bolded font. SD = standard deviation.
biopsy, however, comparison of the patient data to the control group, revealed that some perceptual and acoustic changes had occurred, with some parameters no longer significantly different to the control group performance. As such, it may be suggested that these small shifts in function may be due to the removal of tumour tissue as supported by some authors [14-16].

Examination of vocal change pre and post biopsy, however, revealed no statistically significant difference on the perceptual, acoustic or MPT parameters. On the basis of this finding, it would appear that the biopsy process has minimal or no impact on functional outcomes for early glottic cancer. Although the nature of the biopsy is to remove tissue which could then potentially damage the vocal folds, it is possible in some cases, such as with exophytic tumours, to phonosurgically remove suspect tissue yet preserve the underlying vocal folds as suggested by Melchiors and colleagues [14]. This pilot study, however, did not consistently collect tumour information (exophytic vs. invasive), involvement of the anterior commissure, nor exact volume or type of tissue removed (cancerous versus mucosal, ligamental or muscular tissue), therefore such potential voice impacting factors currently cannot be confirmed. In fact, the results of this study are more congruent with those studies which suggest that type of biopsy do not impact voice outcomes [11,13].

Another possible reason for no significant mean change at a group level may be due to the pilot study’s small sample size and the individual variability within our sample. As revealed by our case by case analysis, half of the group remained unchanged post biopsy on perceptual assessment. However, within the other half, four showed improvements and three showed voice deterioration. Exploration of the patient characteristics for those with improved vocal function, failed to reveal any clear pattern in relation to nature of tumour or biopsy procedure. However, in contrast, the three voice deterioration cases were all punch and biopsied multiple times. This outcome is congruent with authors who have suggested that the punch biopsy may be detrimental to the voice [14] and, furthermore, is an inferior procedure since a single biopsy sample may not be representative, histologically, of the entire pathological tissue [17, 24]. In this study, where patients underwent several punch samples and multiple biopsy procedures, a deterioration to moderate-severe dysphonia was consistently demonstrated. These aforementioned authors and others go on to suggest that the excisional biopsy is the procedure of preference since the punch biopsy may be likely to harm the vocal cords and the excisional biopsy has the dual purpose of being both diagnostic and therapeutic in a single procedure [14, 16, 17, 24]. Claims made by these studies, however, need to be interpreted with caution since specific voice and functional outcomes were not recorded nor rigorously investigated in their study designs.

**Strengths and limitations**

Although this is the first prospective study, to date, investigating voice outcomes pre and post biopsy, several limitations in this pilot study have been identified. This study did not control for types of tumours, and biopsy procedures were not standardized but dependent on surgeon preference, as is reflective of clinical practice. Multiple surgeons across several sites were involved and the operation report (biopsy procedure description) was not always detailed. The small sample size is also a limitation since biopsy technique, tumour size and tumour type as possible factors influencing the post biopsy voice cannot be conclusively suggested with this small sample size. Power calculations (G*Power 3.1.9.2 software [25]), using a priori, 2-tailed, Laplace distribution with a power of 80%, 5% level of significance and effect size of 0.3) suggest that a n=60 would be required in future studies.

However, despite the lack of pre post biopsy group mean statistical difference, the fact that half of the patients showed voice changes post biopsy procedure, suggests that this is an area in need of further research. Future studies should control for types of tumours and biopsy procedures, as far as clinically possible, with a recommended sample size of 60 for adequate power (80%). Pre-post biopsy voice outcomes should be routinely collected in clinical practice.

**Conclusion**

Although it is well established that the biopsy is a necessary first step in laryngeal cancer management, there is emerging discussion and debate surrounding suggested optimal biopsy technique in terms of functional sequelae. In this small cohort, a positive or negative biopsy-voice impact was not confirmed; although individual analyses supported the notion of a negative voice change in those who underwent multiple, punch biopsies. Consequently, this study indicates that type of biopsy, nature and extent of tumour/tissue removal on functional outcomes is an area of further research.

**Authorship and contributorship**

All authors fulfil criteria for authorship according to the International Committee of Medical Journal Editors.

**Funding**


**Conflict of interest**

The authors declare they have no conflicts of interest.