Long-term Voice Outcome Following Radiation *Versus* Laser Microsurgery in Early Glottic Cancer

***Yue Ma, †Rebecca Green, *Stephanie Pan, *Daniel McCabe, *Leanne Goldberg, and *Peak Woo,** **New York, New York, and †New Orleans, Louisiana*

Summary: Objectives. Long-term voice outcome (LTVO) after radiation (XRT) or trans-oral laser microsurgery (TLM) is unclear. This study is a multi-modality analysis of LTVO following XRT or TLM in patients with early glottic cancer. We hypothesize that as compared with TLM, LTVO is worse in the XRT group because of progressive fibrosis in the glottic tissue

Material and Methods. One hundred and two patients with early glottic carcinoma (carcinoma in situ, T1, T2) were included. Multi-modality voice analyses were performed with self-perception using Voice Handicap Index-10, objective analysis using Analysis of Dysphonia in Speech and Voice Software (Cepstral Spectral Index of Dysphonia score for Consensus Auditory-Perceptual Evaluation of Voice sentences), and perceptual rating by two blinded speech language pathologists (GRBAS scale).

Results. Fifty-five patients received TLM (mean follow-up = 52 months) and 47 patients had XRT (mean follow-up = 65 months). There is no difference between the two groups in sex, age, stage, and follow-up time. Intraclass correlation coefficient between raters was high at 0.94. Controlling for age and stage, XRT increases total GRBAS score by 1.38 points ($P = 0.006$) and increases Cepstral Spectral Index of Dysphonia score by 13.7 points ($P < 0.001$) when compared with the TLM group. No significant differences were found in the Voice Handicap Index score between the XRT and the TLM groups.

Conclusions. This is the first multi-modality voice analysis to suggest TLM results in better LTVO than XRT in GRBAS score and objective voice analysis but not in self-perception. These differences may reflect the progressive effects of XRT on glottic tissue. A randomized controlled study is required to confirm our findings.

Key Words: Voice outcome–Laryngeal cancer treatment–Radiation therapy–Laser microsurgery–Quality of life.

INTRODUCTION

Squamous cell carcinoma of the larynx is a common head and neck cancer, with approximately 150,000 new cases diagnosed yearly worldwide.[1–3] Majority of laryngeal cancer arises from the glottis, with 75%–80% presenting at an early stage.[1,2,4,5] Radiation or endolaryngeal excision with or without laser is accepted as single-modality treatments for early-stage glottic cancer (carcinoma in situ [CIS], T1, T2). Radiation therapy (RT) involves a prolonged treatment course with risks of xerostomia, mucositis, and skin changes as well as developing radiation-induced malignancies.[6] Trans-oral laser microsurgery (TLM) gained recent popularity and often represents a desired option as it involves a brief ambulatory surgery course and preserves salvage treatment options for recurrent disease.[7–10]

Because RT and TLM offer similar oncologic outcome, it is important to understand voice quality outcome of each treatment modality.[1,2,6,11–13] Short-term, single-modality analysis of voice outcome from RT or TLM has been investigated in several cohort studies, systematic reviews, and meta-analysis, with varied results.[2,11,13–19] In many studies, only one or two aspects of voice quality are assessed with variable follow-up time. A recent systematic review and meta-analysis suggests no clinical significant difference in Voice Handicap Index (VHI) outcome between TLM and RT for early glottic carcinoma.[19] Most previously published voice results are based on data available within the first 2 years after initiation of treatment. Aaltonen et al. conducted the only randomized control study concerning voice quality after RT or TLM with follow-up time up to 2 years. This study concluded similar overall voice outcome after treatment with RT or TLM.[20] What is not clear from the literature is whether the comparable short-term voice results recorded within the first 2 years between RT and TLM groups can be assumed to be true over the long term (>2 years). Because we expect majority of patients with early glottic carcinoma to have long-term survival, quality of life factors such as voice quality over time between the two groups should be investigated. We conducted a multi-modality analysis of long-term voice outcome (LTVO) following RT or TLM in patients with early glottic cancer. Our study defines LTVO as voice quality at least 2 years beyond the initial treatment. We chose this time point because maturation of TLM wound is generally complete after 18 months. Furthermore, added effects of long-term radiation-related fibrosis would be expected to show progressive changes after the first few years after completion of RT. Such changes may result in voices different from that recorded after initial treatment by TLM or RT. We hypothesize that as compared with TLM, LTVO is worse in the RT group because of progressive fibrosis in the glottic tissue.

Accepted for publication October 25, 2017.

Presented at The American Laryngological Association Annual Meeting 2017, San Diego, CA, USA.

Conflict of interest: There are no financial or personal relationships with other people or organizations that could inappropriately influence (bias) the authors' actions.

Disclosures: There are no specific financial interests, relationship, and affiliations relevant to the subject of the manuscript.

From the *The Icahn School of Medicine at Mt. Sinai, New York, New York; and the †Tulane University School of Medicine, New Orleans, Louisiana.

Address correspondence and reprint requests to Peak Woo, 300 Central Park West, New York, NY 10024. E-mail: Peakwoo@peakwoo.com

Journal of Voice, Vol. 33, No. 2, pp. 176–182

0892-1997

https://doi.org/10.1016/j.jvoice.2017.10.018

TABLE 1.
Study Patient Characteristics

Characteristics	Treatment Groups: Trans-oral Laser Microsurgery (TLM) (n = 55)	Treatment Groups: Radiation Therapy (RT) (n = 47)	*P* Value
Male, n (%)	50/55 (90.9)	40/46 (86.96)	0.525
Age at treatment, mean ± SD, y	65.1 ± 11.4	69.8 ± 12.4	0.052
Age at follow-up mean ± SD, y	69.5 ± 11.6	74.8 ± 12.5	**0.028**
Pathological stage, n (%)			0.100
pTIS	12/55 (21.8)	8/45 (17.8)	
pT1a	32/55 (58.2)	27/45 (60)	
pT1b	10/55 (18.2)	4/45 (8.9)	
pT2	1/55 (1.8)	6/45 (13.3)	
Follow-up time, median (IQR), mo (range)	51.5 (25.8–71.9)	64.7 (36.1–80.0)	0.111

% are expressed as column percentages.
P value < 0.05, bold highlights significance p values <0.05.

MATERIALS AND METHODS

Patients

Between 2009 and 2016, multi-dimensional voice analysis from 125 patients with early glottic cancer (CIS, T1, T2) who underwent either RT or TLM was obtained. All the patients had their cancer treatment completed before 2014, allowing for a minimum of 2 years of posttreatment voice recordings. Patients with previously untreated, biopsy-proven squamous cell carcinomas of the glottis treated with either RT or TLM were identified. Those with at least a 2-year follow-up were included in this study. The multidimensional voice recordings were collected after the completion of treatment as a single measure. No attempt was made to correlate voice before treatment.

The decision between radiation and surgery for early glottis carcinoma was largely by patient preference. The data regarding equivalent survival in early glottic carcinoma between the two options were presented to the patient along with the risks and benefits of treatment. All patients were given a second opinion consultation with a radiation oncologist. Although we cannot rule out some patients were referred to radiation because of poor operative exposure, the near equal distribution between the TLM and the RT groups suggests minimal selection bias by the surgeon.

TLM were performed by a single surgeon (PW). Patients with previous cordectomy, head and neck radiation, or neoadjuvant chemotherapy were excluded. The excluded 23 patients were because of prior treatment by prior radiation in the TLM group or prior TLM in the RT group. Three patients had to have partial laryngectomy. Thus, 102 patients (TLM = 55, RT = 47) who were free of disease, with previously untreated early-stage glottis carcinoma with a minimum of 2 years' follow-up after completion of cancer treatment were included in our final analysis. The characteristics of these patients are reported in Table 1. Within the TLM group, patients were further characterized based on cordectomy type defined by the European Laryngological Society (ELS) classification of endoscopic cordectomies (Table 2).[21,22]

Evaluation of voice quality

Voice analysis was performed with three modalities: self-perception using VHI-10, objective analysis using Analysis of Dysphonia in Speech and Voice (ADSV) Software Cepstral Spectral Index of Dysphonia (CSID) score, and perceptual rating by two blinded speech language pathologists using the GRBAS scale. Speech samples were recorded on the same equipment for both groups. Voice samples from patients treated with either RT or TLM were analyzed; subgroup analysis was subsequently performed to evaluate LTVO treated with different ELS cordectomy types or RT.

VHI-10

The VHI-10 is a psychometrically validated tool developed for measurement of subjective voice disability experienced by patients. Scoring ranges in integer values from 0 to 40. Values ranging from 0 to 10 are considered normal, 11 to 20 are mild, 21 to 30 are moderate, and 31 to 40 are severely impaired.[23]

ANALYSIS OF DYSPHONIA IN SPEECH AND VOICE SOFTWARE CSID SCORE

Running speech samples were collected and analyzed using the ADSV Software (KayPENTAX model 9100B, Montvale, NJ).

TABLE 2.
TLM ELS Cordectomy Types

ELS Classification	TLM n, (%)
I	29/55 (52%)
II	13/55 (24%)
III	0
IV	0
Va	11/55 (20%)
Vb	0
Vc	0
Vd	1/55 (2%)
VI	1/55 (2%)

Patients were asked to repeat the sentence "how hard did he hit him" from the Consensus Auditory-Perceptual Evaluation of Voice sentence four times continuously at regular cadence and loudness for 12 seconds. Subsequently, the software analyzes the recorded token by providing cepstral and spectral acoustic measures that are then combined into a CSID score to estimate dysphonia severity.[24]

PERCEPTUAL ANALYSIS USING THE GRBAS SCALE

Voice samples were uploaded onto a flash drive in a random order. Perceptual analysis of voice quality was performed by two blinded speech pathologists who are experts in voice analysis. Voice quality was evaluated by using the GRBAS scale, consisting of grade (G), roughness (R), breathiness (B), asthenia (A), and strain (S). Ratings of these five aspects of voice quality varied from 0 (normal) to 3 (severely abnormal).[20,25] Inter-rater consistency was assessed by comparing GRBAS scores between the two blinded raters with Cohen kappa and corresponding 95% confidence interval (CI.)

European Laryngological Society classification of endoscopic cordectomies

Operative reports were obtained and reviewed by the surgeon who performed the TLM to identify the type of cordectomy performed. The type of cordectomy was subsequently entered as a variable for subgroup analysis. ELS classification comprises nine types of cordectomies: subepithelial cordectomy (type I); subligamental cordectomy (type II); transmuscular cordectomy (type III); total cordectomy (type IV); extended cordectomy, which encompasses the contralateral vocal fold and the anterior commissure (type Va); extended cordectomy, which includes the arytenoid (type Vb); extended cordectomy, which encompasses the subglottis (type Vc); extended cordectomy, which includes the ventricle (type Vd); and cordectomy (type VI) for cancers arising from the anterior commissure.[21,22]

STATISTICAL ANALYSIS

Patient characteristics are expressed as mean and standard deviations or median and interquartile range (IQR) for continuous variables, frequencies, and percentages for categorical variables. Difference in baseline patient demographics and clinical characteristics were compared using Pearson chi-square tests for categorical variables and Student *t* tests or Mann-Whitney *U* test, when appropriate, for continuous variables. Fisher exact tests were used to compare VHI-10 and severity of dysphonia, in each of the five aspects of the GRBAS scale, between the two treatment groups. Cohen kappa and corresponding 95% CI were used to calculate the inter-rater reliability for GRBAS scale perceptual analysis. A multivariable ordinal logistic regression analysis was adopted to assess the GRBAS score and VHI-10 in each treatment group. Proportional odds assumption was met in all models. Odds ratio and their corresponding 95% CI are reported. A multivariable linear regression analysis was performed to evaluate predictors of CSID scores. All models were adjusted for the patient's age at follow-up and stage of cancer. All *P* values were two-sided, and statistical significance is defined for *P* values <0.05. Statistical analyses were performed using SAS software version 9.4 (SAS Institute, Cary, NC).

RESULTS

In total, 125 patients' charts were reviewed. A total of 102 patients met inclusion criteria and are included in this study, with 55 patients in the TLM and 47 patients in the RT treatment group. Of the 55 patients subjected to TLM, 29 (52%) underwent type I, 13 (24%) type II, 12 (22%) type V, and 1 (2%) type VI cordectomy. The median follow-up time is 51.5 months (IQR, 26–72) for the TLM group and 64.7 months (IQR, 36–80) for the RT group. The mean age at follow-up, when the voice samples were obtained, is 69.5 ± 11.6 for the TLM group and 74.8 ± 12.5 for the RT group. The difference in age at follow-up between the two treatment groups is statistically significant ($P = 0.028$). There is no difference in sex ($P = 0.525$), age at treatment ($P = 0.052$), cancer stage ($P = 0.100$), and median follow-up time ($P = 0.111$) between the two treatment groups. Regression models of treatment effect on voice outcomes are summarized in Table 3.

VHI-10

The mean VHI for the TLM group is 10.69 ± 9.41 and for the RT group is 11.43 ± 8.15. Using VHI-10 values, 29 (53%) patients are classified as normal voice, 17 (31%) as mild, 6 (11%) as moderate, and 3 (5%) as severe dysphonia in the TLM group. In the RT group, VHI-10 values show 24 (57%) with normal voice, 12 (29%) with mild, 5 (12%) moderate, and 1 (2%) with severe dysphonia. Fisher exact test shows no significant difference in the VHI-10 score between the TLM and the RT treatment groups ($P = 0.922$). Table 4 summarizes the VHI-10 results.

CSID score derived from ADSV software

By applying the CSID software to a single repeated sentence "How hard did he hit him" from the Consensus Auditory-Perceptual Evaluation of Voice battery, we were able to derive a single severity of dysphonia score from objective acoustic

TABLE 3.
Regression Models of Treatment Effect on Voice Outcomes in TLM and RT Treatment Groups

	GRBAS Grade (n = 100)		CSID (n = 100)		VHI-10* (n = 97)	
Variables	OR (95% CI)	*P* Value	Estimate (SE)	*P* Value	OR (95% CI)	*P* Value
RT *vs* TLM	3.10 (1.39, 6.92)	**0.006***	13.80 (6.87, 20.73)	**<0.001***	0.84 (0.38, 1.86)	0.661

* Normal: ≤ 10, Mild: 11–20, Moderate: 21–30, Severe: 31–40. References for each variable are as follows: stage: CIS, treatment: TLM. Stage was dichotomized to ensure stable parameter estimates.
P value < 0.05, bold highlights significance p values <0.05.

TABLE 4.
Summary of VHI-10 in TLM and RT Treatment Groups

Variable	Trans-oral Laser Microsurgery				Radiation Therapy (n = 47)†	*P* Value*
	Type I Cordectomy (n = 29)	Type II Cordectomy (n = 13)	Type V Cordectomy (n = 12)	All (n = 55)		
VHI-10	No. of patients (%)					
Normal (≤10)	17 (58.6)	5 (38.5)	7 (58.3)	29 (53)	24 (57)	0.922
Mild (11–20)	8 (27.6)	7 (53.9)	1 (8.3)	17 (31)	12 (29)	
Moderate (21–30)	2 (6.9)	1 (7.7)	3 (25)	6 (11)	5 (12)	
Severe (31–40)	2 (6.9)	0 (0)	1 (8.3)	3 (5)	1 (2)	

* *P* value is calculated for comparisons between all patients on TLM and radiation therapy. % are expressed as column percentages.
† Five patients did not have a VHI-10 score.

measurement of dysphonia (CSID score). Cepstral- and spectral-based measures using CSID to estimate dysphonia severity in continuous speech have demonstrated good reliability. The CSID score is valuable in our patients with moderate dysphonia because unlike prior objective acoustic measures, it can perform the analysis on running speech. The mean CSID score for the TML group is 28.58 ± 14.99 and for the RT group is 42.63 ± 20.12. After controlling for age and stage of disease, there is significant difference in the CSID score between irradiated and laser surgery-treated patients ($P < 0.001$). On average, RT increases CSID score by 13.8 (6.87, 20.73) points as compared with TLM-treated patients.

GRBAS

Cohen kappa between the two speech pathologists is 0.82 (0.72, 0.91), demonstrating high inter-rater reliability. Thus, the scores from the first rater are used for analysis. Voice perception is evaluated by five aspects of GRBAS scale. Table 5 summarizes the severity of dysphonia in the TLM and RT treatment groups. Long-term voice overall grade is statistically different between the two groups ($P = 0.026$), with better LTVO in the TLM group. Controlling for patient age at follow-up and cancer stage, patients treated with RT is 3.1 times (95% CI 1.39, 6.92) more likely to result in a higher overall grade than patients treated with laser surgery ($P = 0.006$). Voice dysfunction in the TLM group tends to be more mild (41.8%) than moderate (38.2%), whereas in the RT group tends to be more moderate (55.3%) than mild (25.5%). There is more normal voice outcome in the TLM group (12.7%) than the RT group (2.1%). More patients had severe dysphonia in the RT group (17%) than TLM group (7.3%). Voice profile in the RT group is considered more rough ($P = 0.020$) and strained ($P = 0.004$). We did not find statistically significant difference in breathiness ($P = 0.649$) and asthenia ($P = 0.367$) between the RT and the TLM groups.

Subgroup analysis

In the subgroup analysis, we compared LTVO in patients treated with type I, type II, or type V cordectomy with patients treated with radiation. As detailed in Table 6, there is no statistically significant difference in VHI between patients who underwent type I ($P = 0.993$), type II ($P = 0.582$), and type V cordectomy ($P = 0.603$), when compared with RT. The mean CSID score for patients who underwent type I cordectomy is 24.45 ± 13.68, for type II is 29.91 ± 15.28, and for type V is 36.32 ± 15.83. There is significant difference in CSID score between patients who underwent type I ($p < 0.001$) or II ($P = 0.011$) resection when compared with RT but not with type V ($P = 0.220$). Perceptual rating by the speech pathologists using the GRBAS scale shows that patients treated with RT is 4.27 times (95% CI 1.62, 11.24,

TABLE 5.
Summary of GRBAS in TLM and RT Treatment Groups*

Variable	Trans-oral Laser Microsurgery (n = 55)	Radiation Therapy (n = 47)	*P* Value
Overall grade			**0.026**
Normal	7 (12.7)	1 (2.1)	
Mild	23 (41.8)	12 (25.5)	
Moderate	21 (38.2)	26 (55.3)	
Severe	4 (7.3)	8 (17.0)	
Roughness			**0.020**
Normal	10 (18.2)	4 (8.5)	
Mild	20 (36.4)	10 (21.3)	
Moderate	22 (40.0)	22 (46.8)	
Severe	3 (5.5)	11 (23.4)	
Breathiness			0.649
Normal	20 (36.4)	18 (38.3)	
Mild	21 (38.2)	15 (31.9)	
Moderate	13 (23.6)	11 (23.4)	
Severe	1 (1.8)	3 (6.4)	
Asthenia			0.367
Normal	37 (67.3)	27 (57.5)	
Mild	14 (25.5)	12 (25.5)	
Moderate	4 (7.3)	6 (12.8)	
Severe	0 (0)	2 (4.3)	
Strain			**0.004**
Normal	23 (41.8)	5 (10.6)	
Mild	11 (20.0)	11 (23.4)	
Moderate	12 (21.8)	19 (40.4)	
Severe	9 (16.4)	12 (25.5)	

* *P* value is calculated for comparisons between all patients on TLM and radiation therapy.
% are expressed as column percentages.
P value < 0.05, bold highlights significance p values <0.05.

TABLE 6.
Regression Models of Cordectomy and Radiation Effect on Voice Outcomes*

	GRBAS Grade (N = 99)		CSID (N = 99)		VHI-10 (N = 96)	
Variables	OR (95% CI)	*P* Value	Estimate (SE)	*P* Value	OR (95% CI)	*P* Value
RT *vs* ELS = I (n = 29)	4.27 (1.62, 11.24)	**0.003**	17.27 (25.49, 9.05)	**<0.001**	1.00 (0.38, 2.62)	0.993
RT *vs* ELS = II (n = 13)	4.99 (1.44, 17.34)	**0.011**	13.93 (24.65, 3.20)	**0.011**	1.40 (0.42, 4.68)	0.582
RT *vs* ELS = V (n = 12)	1.25 (0.36, 4.26)	0.726	6.76 (17.54, 4.03)	0.220	1.38 (0.41, 4.69)	0.603

* Type III analysis *P* value. References for each variable are as follows: stage: CIS, therapy: radiation. Stage was dichotomized to ensure stable parameter estimates.
P value < 0.05, bold highlights significance p values <0.05.

$P = 0.003$) more likely to result in a worse overall voice grade than patients treated with type I cordectomy; 4.99 times (95% CI 1.44, 17.34, $P = 0.011$) when treated with type II cordectomy; and 1.25 times (95% CI 0.36, 4.26, $P = 0.726$) when treated with type V cordectomy. Type I and type II resections did show significant difference in GRBAS scale voice grade when compared with RT, but type V cordectomy did not.

DISCUSSION

Because survival data between TLM and RT treatment are equivalent in patients with early glottic cancer, patients and clinicians frequently ask which modality will give a better functional voice outcome. Voice outcome documented by Aaltonen et al in a randomized control trial suggests overall voice quality up to 2 years after treatment to be similar between RT and TLM groups.[20] However, clinicians often see patients who have had previous RT with complaints of vocal deficiencies secondary to stiffness and chronic laryngitis sicca appearing many years after the initial RT treatment. In the TLM group, there is wound healing and scar formation on the ipsilateral vocal fold that typically concludes in 18 months. On the other hand, in the RT group, progressive fibrosis happens on bilateral vocal cords that may progress beyond 18 months. Although the short-term voice outcome findings may show equivalent results, the LTVO between the TLM and the RT groups may be different. Our study examines LTVO in two similarly matched early glottic cancer groups by sex, age at treatment, cancer stage, and follow-up time treated with either RT or TLM.

Although auditory-perceptual assessment of voice is considered the gold standard of voice evaluation, patient self-reported voice handicap is also important in the process of informed medical decision-making and patient outcome assessment. We did not find any differences between RT and TLM groups based on VHI scores. This may be because of cancer survivors having less concern regarding the quality of their voice. Based on VHI scores alone, this study suggests there are no appreciable patient-perceived differences in voice quality regardless of whether XT or TLM was selected.

However, based on the dimension of objective acoustic analysis, and on quality of voice as perceived by expert voice raters, there may be a difference in voice outcome between treatment type. In recent years, cepstral- and spectral-based measures using CSID to estimate dysphonia severity in continuous speech have demonstrated good reliability.[24,26–29] The CSID-estimated dysphonia severity is theoretically a number between 0 and 100, with 100 being the most severe.[28] Awan et al indicated that for every 1-unit increase in the CSID score, the odds of being in the dysphonia-positive group increases by 10.8%.[27] CSID is also considered a screening tool for voice-disordered cases *versus* normal controls with a cutoff score of 19 (sensitivity of 87% and specificity of 60%).[27] Patients with a CSID score below 19 are considered to have a normal voice. Although CSID is shown to possess a strong relationship with perceptual ratings of dysphonia severity across a variety of vocal pathologies,[26] Awan et al demonstrated that CSID and VHI only correlated at low-to-moderate levels.[28] Therefore, spectral and cepstral acoustic measures and the VHI should be viewed as providing unique and complementary voice outcome information. To capture a robust picture of residual dysphonia severity after treatment, we performed a multi-modality assessment of voice impairment and disability by evaluating self-perception of voice, acoustic parameters of running speech, and perceptual rating from expert listeners who treat dysphonia.

Our study suggests that LTVO may be different from short-term results, which could not differentiate between TLM and RT treatment groups. As a group, this study shows TLM did better in objective measures (CSID scores) and in listener perception scores (GRBAS scores) than the RT group. The mean CSID score for the TLM group demonstrates mild dysphonia, whereas the RT group shows moderate dysphonia based on classification system proposed by Peterson et al (mild: 0–33, moderate: 34–66, and severe: 67–100).[26] Similar outcome was found from perceptual analysis with the GRBAS scale, where TLM LTVO showed more mild impairment when compared with patients treated with RT. Long-term voice profile in patients treated with RT tend to be more rough and strained when compared with patients treated with surgery. No differences in breathiness and asthenia were found in our study. Interestingly, in Aaltonen et al's study, short-term voice quality after TLM was found to be more breathy and weak; however, no differences in grade, roughness, and strain were found.[20] We did not find significant differences in the VHI score between treatment groups. We speculate this finding may be because of our cancer patients being older and having been free of disease, found their voice-related handicaps less significant than the average population.

To evaluate whether voice outcome could be related to the extent of surgical resection, we further divided the TLM group by cordectomy subtypes (Table 2). We did not find significant

TABLE 7.
Regression Models of Cordectomy Effect on Voice Outcomes Among TLM

	GRBAS Grade (n = 55)		CSID (n = 55)		VHI-10* (n = 55)	
Variables	OR (95% CI)	*P* Value	Estimate (SE)	*P* Value	OR (95% CI)	*P* Value
Age at follow-up	1.04 (0.99, 1.09)	0.098	0.07 (-0.25, 0.40)	0.657	1.00 (0.95, 1.05)	0.973
Stage (1a, 1b, 2)	6.28 (1.43, 27.67)	0.015	11.38 (2.00, 20.77)	0.020	3.18 (0.69, 14.52)	0.136
ELS 2+ *vs* 1	1.51 (0.51, 4.42)	0.457	5.41 (-2.10, 12.92)	0.162	1.18 (0.41, 3.46)	0.759

* Normal: ≤ 10, Mild: 11–20, Moderate: 21–30, Severe: 31–40. References for each variable are as follows: stage: CIS, ELS: 1. Stage was dichotomized to ensure stable parameter estimates.

difference in LTVO between patients who were treated with different cordectomy types (Table 7). This is likely because of lack of sufficient number of patients who underwent type III and above resection. Subgroup analysis comparing LTVO in various cordectomy types with radiation demonstrates that type I and II cordectomies have a better LTVO when compared with radiation through auditory-perceptual assessment and objective analysis by CSID. Type V cordectomy did not demonstrate a statically significant difference in LTVO when compared with the RT group. We were surprised that the extent of surgery had relatively minor effect on voice outcome. It could be that the added radiation effect over time on tissue worsens voice in the RT group and masked any effect on the extent of surgery in the type V TLM group. We are not certain why the LTVO results in this study are different from the short-term results. We hypothesize this difference may result from progressive fibrosis of the vocal tract in the radiation group beyond 2 years.

To the best of our knowledge, our study is the largest, single-center, multi-modality analysis of LTVO in early glottic cancer treated by either RT or TLM. There are limitations to our study. Our study is not randomized, which may result in selection bias. However, as the stage between the two treatment groups did not differ, most likely the choice of treatment was made primarily by the patient and not the surgeon. Other factors such as medical morbidity and smoking history between groups were not controlled and may affect posttreatment voice quality; these data were not tabulated for analysis. Our data include only point measures and does not contain serial measures of voice changes with time after treatment. Patients in the RT group are overall older at follow-up; this may contribute to voice differences because of age-related vocal atrophy, independent of their cancer treatment modality. Despite these limitations, the study suggests that LTVO may be different from short-term voice results between TLM and RT. LTVO should be studied more in detail to properly council our patients when deciding on treatment choice in early glottic carcinoma.

CONCLUSION

In summary, our study is the first multi-modality voice analysis that suggests LTVO to be better in TLM than in RT in perceptual and acoustic measures, but not in self-perception VHI-10 score. Voice may change in the long term after cancer treatment in ways that are not yet well studied. Some factors may stem from the long-term effects of radiation on vibratory function of the vocal folds. Our study suggests that, when considering LTVO in patients with early glottic carcinoma, TLM is preferred over RT. A randomized controlled study is required to confirm our findings.

Acknowledgment

The authors thank Ms. Stephanie Pan at Mount Sinai for her valuable assistance with the statistical analysis.

REFERENCES

1. Brady JS, Marchiano E, Kam D, et al. Survival impact of initial therapy in patients with T1-T2 glottic squamous cell carcinoma. *Otolaryngol Head Neck Surg*. 2016;155:257–264.
2. Abdurehim Y, Hua Z, Yasin Y, et al. Transoral laser surgery versus radiotherapy: systematic review and meta-analysis for treatment options of T1a glottic cancer. *Head Neck*. 2012;34:23–33.
3. Thompson L. World Health Organization classification of tumours: pathology and genetics of head and neck tumours. *Ear Nose Throat J*. 2006;85:74.
4. Groome PA, O'Sullivan B, Irish JC, et al. Glottic cancer in Ontario, Canada and the SEER areas of the United States. Do different management philosophies produce different outcome profiles? *J Clin Epidemiol*. 2001;54:301–315.
5. Galletti B, Freni F, Cammaroto G, et al. Vocal outcome after CO2 laser cordectomy performed on patients affected by early glottic carcinoma. *J Voice*. 2012;26:801–805.
6. Warner L, Chudasama J, Kelly CG, et al. Radiotherapy versus open surgery versus endolaryngeal surgery (with or without laser) for early laryngeal squamous cell cancer. *Cochrane Database Syst Rev*. 2014;(12):CD002027.
7. Ansarin M, Cattaneo A, De Benedetto L, et al. Retrospective analysis of factors influencing oncologic outcome in 590 patients with early-intermediate glottic cancer treated by transoral laser microsurgery. *Head Neck*. 2017;39:71–81.
8. Wilkie MD, Lightbody KA, Lythgoe D, et al. Transoral laser microsurgery for early and moderately advanced laryngeal cancers: outcomes from a single centralised United Kingdom centre. *Eur Arch Otorhinolaryngol*. 2015;272:695–704.
9. Goor KM, Peeters AJ, Mahieu HF, et al. Cordectomy by CO2 laser or radiotherapy for small T1a glottic carcinomas: costs, local control, survival, quality of life, and voice quality. *Head Neck*. 2007;29:128–136.
10. Diaz-de-Cerio P, Preciado J, Santaolalla F, et al. Cost-minimisation and cost-effectiveness analysis comparing transoral CO(2) laser cordectomy, laryngofissure cordectomy and radiotherapy for the treatment of T1-2, N0, M0 glottic carcinoma. *Eur Arch Otorhinolaryngol*. 2013;270:1181–1188.
11. Kerr P, Mark Taylor S, Rigby M, et al. Oncologic and voice outcomes after treatment of early glottic cancer: transoral laser microsurgery versus radiotherapy. *J Otolaryngol Head Neck Surg*. 2012;41:381–388.
12. Higgins KM, Shah MD, Ogaick MJ, et al. Treatment of early-stage glottic cancer: meta-analysis comparison of laser excision versus radiotherapy. *J Otolaryngol Head Neck Surg*. 2009;38:603–612.
13. Feng Y, Wang B, Wen S. Laser surgery versus radiotherapy for T1-T2N0 glottic cancer: a meta-analysis. *ORL J Otorhinolaryngol Relat Spec*. 2011;73:336–342.

14. van Loon Y, Sjogren EV, Langeveld TP, et al. Functional outcomes after radiotherapy or laser surgery in early glottic carcinoma: a systematic review. *Head Neck*. 2012;34:1179–1189.
15. van Gogh CD, Verdonck-de Leeuw IM, Wedler-Peeters J, et al. Prospective evaluation of voice outcome during the first two years in male patients treated by radiotherapy or laser surgery for T1a glottic carcinoma. *Eur Arch Otorhinolaryngol*. 2012;269:1647–1652.
16. Spielmann PM, Majumdar S, Morton RP. Quality of life and functional outcomes in the management of early glottic carcinoma: a systematic review of studies comparing radiotherapy and transoral laser microsurgery. *Clin Otolaryngol*. 2010;35:373–382.
17. Sjogren EV, van Rossum MA, Langeveld TP, et al. Voice outcome in T1a midcord glottic carcinoma: laser surgery vs radiotherapy. *Arch Otolaryngol Head Neck Surg*. 2008;134:965–972.
18. Osborn HA, Hu A, Venkatesan V, et al. Comparison of endoscopic laser resection versus radiation therapy for the treatment of early glottic carcinoma. *J Otolaryngol Head Neck Surg*. 2011;40:200–204.
19. Greulich MT, Parker NP, Lee P, et al. Voice outcomes following radiation versus laser microsurgery for T1 glottic carcinoma: systematic review and meta-analysis. *Otolaryngol Head Neck Surg*. 2015;152:811–819.
20. Aaltonen LM, Rautiainen N, Sellman J, et al. Voice quality after treatment of early vocal cord cancer: a randomized trial comparing laser surgery with radiation therapy. *Int J Radiat Oncol Biol Phys*. 2014;90:255–260.
21. Remacle M, Eckel HE, Antonelli A, et al. Endoscopic cordectomy. A proposal for a classification by the Working Committee, European Laryngological Society. *Eur Arch Otorhinolaryngol*. 2000;257:227–231.
22. Remacle M, Van Haverbeke C, Eckel H, et al. Proposal for revision of the European Laryngological Society classification of endoscopic cordectomies. *Eur Arch Otorhinolaryngol*. 2007;264:499–504.
23. Rosen CA, Lee AS, Osborne J, et al. Development and validation of the voice handicap index-10. *Laryngoscope*. 2004;114:1549–1556.
24. Awan SN, Roy N, Jette ME, et al. Quantifying dysphonia severity using a spectral/cepstral-based acoustic index: comparisons with auditory-perceptual judgements from the CAPE-V. *Clin Linguist Phon*. 2010;24:742–758.
25. Isshiki N, Okamura H, Tanabe M, et al. Differential diagnosis of hoarseness. *Folia Phoniatr (Basel)*. 1969;21:9–19.
26. Peterson EA, Roy N, Awan SN, et al. Toward validation of the cepstral spectral index of dysphonia (CSID) as an objective treatment outcomes measure. *J Voice*. 2013;27:401–410.
27. Awan SN, Roy N, Zhang D, et al. Validation of the Cepstral Spectral Index of Dysphonia (CSID) as a screening tool for voice disorders: development of clinical cutoff scores. *J Voice*. 2016;30:130–144.
28. Awan SN, Roy N, Cohen SM. Exploring the relationship between spectral and cepstral measures of voice and the Voice Handicap Index (VHI). *J Voice*. 2014;28:430–439.
29. Awan SN, Roy N, Dromey C. Estimating dysphonia severity in continuous speech: application of a multi-parameter spectral/cepstral model. *Clin Linguist Phon*. 2009;23:825–841.