Voice Outcomes Following Radiation Versus Laser Microsurgery for T1 Glottic Carcinoma: Systematic Review and Meta-analysis

Matthew T. Greulich, MD¹, Noah P. Parker, MD²,³, Philip Lee, MD¹, Albert L. Merati, MD⁴, and Stephanie Misono, MD, MPH¹

Abstract

Objective. Systematic review of literature on patient-reported voice handicap following T1 glottic squamous cell carcinoma treatment using transoral laser microsurgery or radiation therapy.


Review Methods. These data sources were searched for papers reporting Voice Handicap Index (VHI) after treatment of early glottic carcinoma. Review and reference cross-checking were performed using a priori selection criteria. Study data were abstracted and publication quality categorized independently by 2 authors. Corresponding authors were contacted to maximize data for analysis. Meta-analysis was performed only with studies that included both treatment modalities, to reduce heterogeneity and maximize rigor; random effects modeling was used to pool results.

Results. Eighteen publications were identified that reported VHI data following surgery and radiotherapy for T1 glottic carcinoma. No studies were randomized. When studies that reported multiple T-stages or systematic treatment selection bias were excluded, 8 retrospective cohort studies describing 362 patients were suitable for meta-analysis. Follow-up time (mean, 47 months; range, 1-298 mo) and extent of surgical excision varied across studies. Six studies showed no VHI difference between treatment arms; 2 favored radiotherapy over surgery (1 of which reported transmuscular cordectomy for all surgical patients); and none favored surgery. Meta-analysis showed no significant difference in post-treatment VHI between radiotherapy and surgery (mean difference, −5.52; 95% confidence interval, −11.40, 0.36; heterogeneity $I^2 = 61\%$, $P = .01$).

Conclusion. VHI scores were comparable following transoral laser microsurgery and radiation therapy for T1 glottic carcinoma in the current literature, suggesting no clinically significant difference in functional voice outcomes between treatment types.

Keywords

humans, laryngeal neoplasms, laser therapy, neoplasm staging, glottis/radiotherapy, glottis/surgery, voice quality, voice disorders, quality of life

Received October 25, 2014; revised January 23, 2015; accepted February 20, 2015.

Squamous cell carcinoma of the head and neck arises most commonly in the larynx, and approximately 75% of cases involve the glottis. Most glottic squamous cell carcinoma cases (75%-80%) present at an early stage² and may be treated by surgery or radiation therapy (XRT) as a single modality. While open surgical procedures are an option, transoral microsurgery has become the technique of choice; this may utilize cold techniques or a variety of lasers (most commonly, the carbon dioxide laser). XRT and transoral laser microsurgery (TLM) are generally thought to have comparable oncologic and survival outcomes. Therefore, decisions regarding use of XRT versus TLM have focused on other factors, including quality-of-life outcomes, particularly those related to vocal function.

In a 2006 publication, Cohen et al⁴ compared Voice Handicap Index (VHI) outcomes of treatment of early glottic carcinoma with XRT and TLM using a fixed effects model meta-analysis of 6 studies. This included 4 case series of patients treated with either XRT alone or TLM alone; there

¹Department of Otolaryngology, University of Minnesota Medical Center, Minneapolis, Minnesota, USA
²The Voice Clinic of Indiana, Carmel, Indiana, USA
³The Department of Speech and Hearing Sciences, Indiana University, Bloomington, Indiana, USA
⁴Department of Otolaryngology, University of Washington Medical Center, Seattle, Washington, USA

This article was presented at the 2014 AAO-HNSF Annual Meeting & OTO EXPO; September 21-24, 2014; Orlando, Florida.

Corresponding Author:
Stephanie Misono, Department of Otolaryngology, University of Minnesota, 420 Delaware St SE, Minneapolis, MN 55455, USA.
Email: smisono@umn.edu
were few comparative studies available for inclusion at that time. More recently, Higgins et al. performed a literature review, with meta-analysis comparing objective voice outcomes, videostroboscopy findings, and GRBAS scoring across treatment types; comparative studies and single-treatment arm studies were included, and VHI data were not among the outcomes.

The objective of this study was to perform an updated systematic review and meta-analysis of voice outcome data for T1 glottic carcinoma given the more recent publication of several papers reporting VHI outcomes for both treatment types. To reduce clinical heterogeneity and maintain methodological rigor, this study focused only on the results of patients with T1 tumors and limited meta-analysis to comparative papers that reported outcomes from both treatment types.

Materials and Methods

Outcome Measure: Voice Handicap Index

The VHI is a patient self-reported instrument in broad use; it has been translated into (and validated for use in) many languages. The VHI consists of 30 questions spanning functional, physical, and emotional subdomains. The total score ranges from 0 to 120, with a lower score indicating a less severe patient-reported voice-related handicap. The VHI-10 is a shorter version consisting of 10 questions selected from the original VHI, with a scoring range of 0 to 40. This has similarly been shown to be valid and responsive to treatment in patients with laryngeal pathology. Papers reporting VHI-10 results were normalized to the standard VHI by multiplying the reported mean and standard deviation by the median (0.411, equating to a multiplicative factor of 1/0.411, or 2.43) of previously published conversion ratios. Sensitivity analysis was performed using the upper and lower extremes of the range of conversion ratios presented for a variety of pathologies; this did not change the findings of the meta-analysis. Given the international nature of this review, we also confirmed that translation of the VHI into each language involved had been assessed for validity.

Literature Search

A systematic literature search using PubMed, Medline, EMBASE, Scopus, Web of Science, and CINAHL was undertaken for papers available in the English language between 1997, when the VHI was published, and 2013. All papers were selected for abstract review that described voice outcomes as measured by the VHI in patients with early-stage glottic or laryngeal squamous cell carcinoma. The search strategy used is shown in Figure 1. The search results were pooled and duplicate results removed. References of papers meeting criteria for systematic review were also cross checked to identify additional relevant papers. The study was exempt from Institutional Review Board review, as it used existing literature.

Systematic Review

Inclusion criteria for systematic review were defined as those papers comparing VHI outcomes for adult patients who had undergone primary TLM versus XRT for early glottic cancers. Abstract review was performed by 3 authors (M.T.G., N.P.P., P.L.) for inclusion in the systematic review. If title and abstract review did not provide sufficient data, the full text of the paper was screened for inclusion criteria.

Full-text review of each paper meeting the above inclusion criteria was independently performed by 2 authors using a standardized data abstraction sheet. Each author assessed the quality of the studies, and differences between reports were resolved by discussion. These quality judgments were summarized for each paper. Additionally the Newcastle-Ottawa Assessment Scale was used to systematically assess study quality.

Meta-analysis

Inclusion in the meta-analysis required adequate reporting of the mean and standard deviation for XRT and TLM groups and an ability to differentiate the VHI results of T1 glottic carcinoma patients from those with Tis or more advanced cancer stages. Corresponding authors were contacted as needed to distinguish T1 stage tumor outcomes from other stages or to acquire additional data needed for meta-analysis. Papers with explicitly stated selection bias were excluded from the meta-analysis.

Statistical Methods

Meta-analysis was performed using Review Manager software available from the Cochrane Collaboration’s website (RevMan 5.2). A random effects model was used given the heterogeneity inherent in the literature. Mean difference was used for statistical comparison to allow assessment of the clinical significance of the results obtained—that is, whether the results were within the test-retest variability of the VHI.
In cases where VHI median and range were not available, mean and standard deviation were estimated using the method described by Hozo et al.16 As a secondary analysis, the effects of more extensive surgical resection on VHI outcomes were also reviewed. VHI outcomes following class ELS17,18 (European Laryngological Society) I&II resection were compared to those following ELS III resection. Data were derived from papers included in the systematic review and several separate studies identified during the literature review process, including single-arm studies that reported only on post-TLM outcomes. If the data were not reported as being grouped between ELS I&II and ELS III, the mean and pooled standard deviation using Bessel’s correction were calculated from the available reported data. A Student’s t test was performed comparing VHI outcomes for ELS I&II and ELS III for those papers with enough information to perform this statistical test. Those without adequate information to allow comparison were considered not applicable.

Results

Literature Search

The initial literature search returned 374 results, 118 of which were unique, nonduplicate papers and abstracts. Reference cross-checking identified 82 additional papers of potential relevance for a total of 200 papers. Abstract and content review identified 18 papers meeting criteria for inclusion in the systematic review. All papers included were cohort studies comparing XRT with TLM excision; no randomized controlled trials were found. Several conference abstracts were also discovered in the computerized search, but none contained sufficient data for inclusion. Reasons for study exclusion are summarized in Figure 2.

Systematic Review of Literature

Ten of the 18 papers selected for systematic review were excluded from the meta-analysis—5 for explicitly stated treatment selection bias, 1 due to incomplete VHI results, and 4 due to insufficient data for meta-analysis.19-28 These papers are summarized in Table 1, as they report important and data not represented in the results of meta-analysis. Two studies19,22 reported using the presence of intact mucosal wave on videostroboscopy to determine eligibility for TLM treatment and found VHI outcomes favoring TLM. The remainder of these relevant but excluded papers observed no significant difference (2 papers) or reported results favoring XRT (3 papers, of which 2 presented results in which T1 data could not be distinguished from those of other stages). Three additional papers did not allow us to distinguish between XRT and TLM results, as they were presented as a combined group in the context of other comparisons.26-28

Meta-analysis: Study Quality and Results

Aspects of study quality for the 8 papers meeting inclusion criteria for meta-analysis are reviewed in Table 2.29-36 Sources of possible patient selection bias included the lack of exclusion of patients with recurrence disease and/or the need for second treatment with either surgery or XRT in 3 studies.29,31,34 Consecutive enrollment was carried out in 3 of the 8 papers, and 1 additional paper discussed prospective collection of voice outcome data. No studies incorporated randomization; most stated that treatment decisions were made by patient preference,29,34-36 and 3 did not specify treatment selection methods.30-32 The study by Sjogren et al33 included patients staged T1a, and all patients were considered eligible for both treatment options with a historical treatment selection method; all patients after 1999 received TLM, and all patients prior to 1996 received XRT. Administration route of the VHI was variable. In many cases, it was difficult to determine the time from treatment to follow-up VHI assessment, but all papers had mean/median follow-up times >6 months. None of the included papers presented baseline VHI data or the use of voice or other therapy. Comorbidity data were not available.

A total of 207 patients receiving XRT and 155 patients receiving transoral surgical excision were included in the meta-analysis. Demographic, treatment, and follow-up data are summarized in Table 3. Most demonstrated similar age distribution and male preponderance with variable follow-up duration and stage (T1a and T1b) distribution. VHI results from the studies are listed in Figure 3. Two of the included studies29,31 showed a significant result favoring XRT over surgical intervention. The remaining 6 papers
showed nonsignificant differences between treatment modalities. Meta-analysis of the VHI outcomes slightly favored XRT over TLM resection but did not reach significance (mean difference, -5.52; 95% CI, –11.40, 0.36). Heterogeneity testing found significant heterogeneity in the data (\(\chi^2 = 17.79, P = .01, I^2 = 61\%\)). Funnel plot showed no evidence for publication bias (not shown, available upon request).

**Effect of Extent of Resection on VHI Outcomes**

ELS classification of resection was available in 6 papers, including 2 that were included in the meta-analysis above and 4 identified during the systematic review but excluded from meta-analysis due to lack of a comparative XRT arm. While a majority of early-stage glottic lesions underwent ELS I or II resections, a large proportion of patients with known ELS classification underwent type III or greater resection. Generally, better VHI outcomes were reported following lower- (ELS I&II) versus higher-grade resections (ELS III), although none of these differences was statistically significant (Table 4).31,36-40

**Discussion**

Results of this meta-analysis suggest no significant difference between subjective voice outcomes following XRT versus surgery for treatment of T1 glottic carcinoma, with a trend toward slightly better scores in the XRT group. The quantitative difference between the XRT and TLM groups was within the test-retest reliability of the VHI.6 Thus, these data suggest that differences in voice-related handicap outcomes between XRT and TLM are clinically negligible. These findings are in keeping with prior literature identifying comparable and relatively mild voice handicap following both treatment types.4,5,41-43

A major strength of this meta-analysis is the inclusion of more primary papers than in prior studies due to the increased number of eligible publications in the literature as well as the willingness of corresponding authors to provide additional information as needed. Rigorous methodology was also used to minimize bias.

Despite these strengths, there remain caveats to the findings presented here. Most important, because treatment type was not randomly selected, treatment allocation is a source of probable bias. However, in the absence of randomized studies, this approach allows us to learn as much as possible from the existing literature in a systematic and consistent fashion. The subjective nature of the VHI also lends itself to the possibility of variability among patients and could be influenced by patient expectations and pretreatment counseling, which may reduce its sensitivity to more subtle changes in voice outcomes. We did not identify any papers that presented pretreatment VHI as a potentially important determinant of posttreatment VHI, and little information was provided on this topic.

<table>
<thead>
<tr>
<th>Author</th>
<th>Reason for Exclusion</th>
<th>XRT n</th>
<th>VHI±</th>
<th>TLM n</th>
<th>VHI±</th>
<th>Reported Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goor et al19</td>
<td>Reported treatment selection bias</td>
<td>23</td>
<td>17.1 ± NR</td>
<td>42</td>
<td>10.6 ± NR</td>
<td>No additional analysis</td>
</tr>
<tr>
<td>Peeters et al22</td>
<td>Reported treatment selection bias</td>
<td>46</td>
<td>18 ± NR</td>
<td>56</td>
<td>12 ± NR</td>
<td>Significant result favoring TLM</td>
</tr>
<tr>
<td>Siupsinskiene et al25</td>
<td>Unable to exclude data for sites outside glottis</td>
<td>20</td>
<td>38.5 ± 26.2</td>
<td>12</td>
<td>41.58 ± 30.1</td>
<td>NS</td>
</tr>
<tr>
<td>Kujath et al20</td>
<td>Unable to exclude non-T1 stage data</td>
<td>25</td>
<td>NR</td>
<td>54</td>
<td>NR</td>
<td>TLM significantly more likely to have VHI &gt; 10</td>
</tr>
<tr>
<td>Kerr et al23</td>
<td>Unable to exclude non-T1 stage data</td>
<td>49</td>
<td>NR</td>
<td>83</td>
<td>NR</td>
<td>Reports median VHI favoring XRT</td>
</tr>
<tr>
<td>Remmelts et al24</td>
<td>Incomplete VHI data (physical subdomain only)</td>
<td>45</td>
<td>8.3 ± 7.7</td>
<td>44</td>
<td>12.4 ± 8.9</td>
<td>NS in T1a Significant result favoring XRT in T1b</td>
</tr>
<tr>
<td>Mlynarek et al21</td>
<td>Unable to exclude non-T1 stage data</td>
<td>4</td>
<td>52.75 ± NR</td>
<td>5</td>
<td>44.8 ± NR</td>
<td>NS</td>
</tr>
<tr>
<td>Van Gogh et al26</td>
<td>Reported treatment selection bias</td>
<td>126</td>
<td>NR</td>
<td>51</td>
<td>NR</td>
<td>XRT and TLM not compared</td>
</tr>
<tr>
<td>Van Gogh et al27</td>
<td>Reported treatment selection bias</td>
<td>126</td>
<td>NR</td>
<td>51</td>
<td>NR</td>
<td>XRT and TLM not compared</td>
</tr>
<tr>
<td>Van Gogh et al28</td>
<td>Reported treatment selection bias</td>
<td>24</td>
<td>NR</td>
<td>11</td>
<td>NR</td>
<td>XRT and TLM not compared</td>
</tr>
</tbody>
</table>

**Table 1. Excluded Studies.**

Abbreviations: NR, not reported; NS, nonsignificant; TLM, transoral laser microsurgery; XRT, radiation therapy.

±Mean ± standard deviation.

814 Otolaryngology–Head and Neck Surgery 152(5)

Downloaded from oto.sagepub.com at SOCIEDADE BRASILEIRA DE CIRURGIA on May 11, 2015
<table>
<thead>
<tr>
<th>Author</th>
<th>Patient Selection</th>
<th>Treatment Selection</th>
<th>Technique Reporting</th>
<th>VHI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dinapoli et al&lt;sup&gt;29&lt;/sup&gt;</td>
<td>No exclusion criteria given</td>
<td>Patient preference, except anatomic / medical reasons</td>
<td>Not specified</td>
<td>CO2 laser</td>
</tr>
<tr>
<td>Loughran et al&lt;sup&gt;30&lt;/sup&gt;</td>
<td>Excluded multiple modalities or recurrence</td>
<td>NR</td>
<td>Conventional</td>
<td>CO2 laser</td>
</tr>
<tr>
<td>Luo et al&lt;sup&gt;35&lt;/sup&gt;</td>
<td>Consecutive enrollment Excluded multiple modalities or recurrence</td>
<td>Patient preference except anatomic / medical reasons</td>
<td>Conventional</td>
<td>CO2 laser, ELS class</td>
</tr>
<tr>
<td>Nunez Batalla et al&lt;sup&gt;31&lt;/sup&gt;</td>
<td>All patients free of disease at time of study No additional exclusion criteria</td>
<td>NR</td>
<td>Not specified</td>
<td>CO2 laser, ELS class</td>
</tr>
<tr>
<td>Oridate et al&lt;sup&gt;32&lt;/sup&gt;</td>
<td>Definitive therapy only Excluded recurrence</td>
<td>NR</td>
<td>Conventional</td>
<td>Laser type not specified, ELS class</td>
</tr>
<tr>
<td>Sjogren et al&lt;sup&gt;33&lt;/sup&gt;</td>
<td>Consecutive enrollment Excluded patients with anterior commissure involvement Excluded multiple modalities or recurrence Excluded patients not eligible for both treatment options</td>
<td>Year of treatment, historical radiation treatment group</td>
<td>Conventional</td>
<td>Laser type not specified, ELS class</td>
</tr>
<tr>
<td>Tomifufi et al&lt;sup&gt;36&lt;/sup&gt;</td>
<td>Consecutive enrollment Excluded recurrence after author communication</td>
<td>Patient preference</td>
<td>Not specified</td>
<td>CO2 laser, ELS class</td>
</tr>
<tr>
<td>Taylor et al&lt;sup&gt;34&lt;/sup&gt;</td>
<td>Prospective patient enrollment All patients T1b</td>
<td>Patient preference based on tumor board recommendation</td>
<td>Not specified</td>
<td>Laser type not specified</td>
</tr>
</tbody>
</table>

Abbreviations: ELS, European Laryngological Society; NR, not reported; NS, not specified; Pts, patients; VHI, Voice Handicap Index.

<sup>a</sup>Shaded cells indicate possible sources of bias.
available on the use of speech therapy. There may also be some cultural differences that influenced VHI scores across studies, although this would not be likely to account for differences between treatment groups within any given study. Conceptually, more extensive vocal fold resections would be thought to worsen voice, and this analysis was included in this review given the high proportion of patients with T1 tumors undergoing ELS III or greater resections in the included studies. However, the current literature does not support such a correlation. This may be due to a true lack of significant difference, as has been proposed, or may reflect a lack of statistical power to detect worse voice quality associated with greater resection.

Table 3. Patient Demographics Summary.

<table>
<thead>
<tr>
<th>Study</th>
<th>Mean Age, y</th>
<th>M/F Ratio</th>
<th>Mean</th>
<th>Range</th>
<th>XRT Dose, Gy</th>
<th>Tla</th>
<th>Tlb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dinapoli et al29</td>
<td>64*</td>
<td>14.9</td>
<td>60</td>
<td>N/A</td>
<td>70</td>
<td>48</td>
<td>9 XRT</td>
</tr>
<tr>
<td>Loughran et al30</td>
<td>69.8</td>
<td>4.5</td>
<td>31.4</td>
<td>XRT</td>
<td>N/A</td>
<td>55</td>
<td>18 XRT</td>
</tr>
<tr>
<td>Luo et al35</td>
<td>68</td>
<td>20</td>
<td>28.4</td>
<td>XRT</td>
<td>16-69</td>
<td>65-70</td>
<td>27.6 TLM</td>
</tr>
<tr>
<td>Nunez Batalla et al31</td>
<td>65.5</td>
<td>N/A</td>
<td>43</td>
<td>6-81</td>
<td>65</td>
<td>11</td>
<td>9 XRT</td>
</tr>
<tr>
<td>Oridate et al32</td>
<td>70</td>
<td>12.7</td>
<td>38</td>
<td>1-298</td>
<td>65</td>
<td>32</td>
<td>11 XRT</td>
</tr>
<tr>
<td>Sjogren et al33</td>
<td>67.9</td>
<td>4.8</td>
<td>60</td>
<td>XRT</td>
<td>15-82</td>
<td>N/A</td>
<td>16 XRT</td>
</tr>
<tr>
<td>Taylor et al34</td>
<td>66.7</td>
<td>9.5</td>
<td>34</td>
<td>5-102</td>
<td>N/A</td>
<td>0</td>
<td>13 XRT</td>
</tr>
<tr>
<td>Tomifuji et al36</td>
<td>68.1</td>
<td>23.5</td>
<td>N/A</td>
<td>6-12</td>
<td>60-70</td>
<td>17</td>
<td>16 XRT</td>
</tr>
</tbody>
</table>

Abbreviations: N/A, not available; M/F, male/female; TLM, transoral laser microsurgery; XRT, radiation therapy.

*Median.

Figure 3. Meta-analysis results and forest plot. CI, confidence interval; IV, inverse variance; TLM, transoral laser microsurgery; VHI, Voice Handicap Index; XRT, radiation therapy.

The choice between treatment types requires consideration of a number of factors. Given that survival is similar between radiation and surgery (if not slightly better following surgery), and that functional voice outcomes appear to be comparable, other characteristics need to be taken into account. While voice outcomes may be equivalent, issues of cost, duration of treatment (and time away from work), and organ preservation more consistently favor surgery in the current literature. Taken together, these factors suggest that surgical excision should be given serious consideration when selecting treatment for T1 glottic carcinoma, although a careful discussion of the uncertainties, risks, and benefits of each method of treatment with the patient remains the cornerstone in...
treatment decision making for this disease. Limitations of the existing literature necessitate randomized studies comparing outcomes across different treatment types to add crucial additional information to this decision-making process.

**Conclusions**

VHI scores were comparable following XRT and TLM for T1 glottic carcinoma in the current literature, suggesting no clinically significant difference in functional voice outcomes between treatment types. No randomized studies were identified, highlighting a need for further investigation in this area.

**Acknowledgments**

We gratefully acknowledge the assistance of Judith Stanke, Biomedical Library, University of Minnesota Medical School, for designing the search strategies for this study. We also appreciate the gracious communication of authors whom we contacted for additional information as needed.

**Author Contributions**

Matthew T. Greulich, data acquisition, analysis, and interpretation; draft and critical revision of manuscript; final approval; Noah P. Parker, data acquisition, analysis, and interpretation; draft and critical revision of manuscript; final approval; Phillip Lee, data acquisition and analysis; critical revision of manuscript; final approval; Albert L. Merati, study conception; data interpretation; critical revision of manuscript; final approval; Stephanie Misono, study conception, design; data interpretation; draft and critical revision of manuscript; final approval.

**Disclosures**

**Competing interests:** None.

**Sponsorships:** None.

**Funding source:** None.

**References**


Handicap Index (VHI-30) and its shortened version (VHI-10)]. Acta Otorhinolaringol Esp. 2007;58:386-392.


