

## Voice in early glottic cancer compared to benign voice pathology

C. D. L. Van Gogh · H. F. Mahieu · D. J. Kuik ·  
R. N. P. M. Rinkel · J. A. Langendijk ·  
I. M. Verdonck-de Leeuw

Received: 7 August 2006 / Accepted: 3 April 2007 / Published online: 25 April 2007  
© Springer-Verlag 2007

**Abstract** The purpose of this study is to compare (Dutch) Voice Handicap Index (VHI<sub>vumc</sub>) scores from a selected group of patients with voice problems after treatment for early glottic cancer with patients with benign voice disorders and subjects from the normal population. The study included a group of 35 patients with voice problems after treatment for early glottic cancer and a group of 197 patients with benign voice disorders. Furthermore, VHI scores were collected from 123 subjects randomly chosen from the normal population. VHI reliability was high with high internal consistency and test–retest stability. VHI scores of glottic cancer patients were similar to those of patients with voice problems due to benign lesions. Both groups of patients were clearly deviant from the normal population. Within the normal population, 16% appeared to have not-normal voices. Based on ROC curves a cut-off score of 15 points was defined to identify patients with voice problems in daily life. A clinical relevant difference score of 10 points was defined to be used for individual patients and of 15 points to be used in study designs with groups. Patients with voice problems after treatment for early glottic cancer encounter

the same amount of problems in daily life as the other voice-impaired patients. The VHI proved to be an adequate tool for baseline and effectiveness measurement of voice.

**Keywords** Voice handicap index · Quality of life · Norm values · Early glottic cancer · Voice impairment

### Introduction

Voice impairment in patients after treatment for early glottic cancer has been reported in several studies, ranging from 14 to 92% of the patients [1, 6, 9, 16, 24, 25]. Furthermore, several studies on the influence of voice problems on quality of life revealed that in 27 up to 58% of the patients experienced difficulties in communication abilities leading to a disrupted social life [3, 8, 12, 17, 19–22, 25]. To enable quick screening on voice problems, a short 5-item voice-screening questionnaire was developed and validated, which proved to be feasible in clinical practice [22]. A more detailed multidimensional voice analysis protocol is however recommended for monitoring voice intervention and for research purposes [24], including a structured questionnaire such as the Voice handicap index (VHI). The VHI is a validated 30-item questionnaire measuring psychosocial handicapping effects of voice disorders [7] and is used in several studies on patients after treatment for early glottic cancer with mean VHI scores ranging from 12 to 34 points [3, 8, 12, 19]. Most of these studies include patients with and without deviant voice quality and mean VHI data are therefore not informative on the amount of problems that patients with voice impairment after oncological treatment encounter in daily life. In a study on 23 patients with voice problems after oncological treatment, Van Gogh et al. [23] reported a mean VHI score of 35. However,

---

C. D. L. Van Gogh · H. F. Mahieu · R. N. P. M. Rinkel ·  
I. M. Verdonck-de Leeuw (✉)  
Department of Otolaryngology Head and Neck Surgery,  
VU University Medical Centre, P.O. box 7057,  
1007 MB Amsterdam, The Netherlands  
e-mail: im.verdonck@vumc.nl

D. J. Kuik  
Department of Clinical Epidemiology and Biostatistics,  
VU University Medical Center,  
Amsterdam, The Netherlands

J. A. Langendijk  
Department of Radiation Oncology,  
University Medical Center Groningen,  
Groningen, The Netherlands

interpretation of how cancer patients cope with voice problems compared to patients with voice problems due to benign laryngeal lesions and compared with the normal population is difficult because of some underexposed psychometric characteristics of the VHI: data from the normal population are limited, no clear clinical cut-off score is available, and information on clinical relevant difference scores is scarce.

The purpose of this study is to compare voice problems of patients after treatment for early glottic cancer with voice problems as reported by patients with benign voice disorders and subjects from the normal population. The study will also provide psychometric information of the VHI regarding internal consistency, reliability, normative data and a clinical cut-off score, and clinical relevant difference scores for use in individual patients and group study designs.

## Materials and methods

### Patients

The patient sample consisted of 232 subjects: 35 patients with voice problems after treatment for early glottic cancer and 197 patients with voice problems due to benign voice disorders.

Patients after treatment for early glottic cancer (carcinoma in situ, T1 and T2 tumours) were selected based on a validated voice-screening questionnaire; having a voice problem was defined as a score of 5 or higher (on a 10-point scale) on one of the 5 voice items [22]. Of these 35 patients, 33 were males, 2 females; the median age was 62 years (range: 41–81); mean post-oncological treatment time was 32 months (range: 6–135). Treatment included radiotherapy ( $n = 24$ ) or endoscopic laser surgery ( $n = 11$ ); mean VHI scores regarding treatment modality were comparable (37 vs. 36 points).

Patients with voice problems due to benign voice disorders were randomly selected from the patient population at our voice clinic. This cohort of 197 patients included 44 patients with vocal fold paresis, 84 with structural lesions (polyps, nodules, scarring, granuloma), 10 patients with Reincke's oedema, 55 patients with laryngitis, and 5 patients with laryngeal trauma. Of these 197 patients, 82 were males and 115 females; median age was 46 years (range 18–90).

### Controls

The group of 123 randomly selected controls from the normal population (employees from the hospital and (acquaintances of) relatives and neighbours of the researchers)

consisted of 54 males and 58 females (gender was not indicated by 11 subjects); median age was 55 years (range 23–87).

### Voice handicap index

The VHI is a validated questionnaire measuring psychosocial handicapping effects of voice disorders and was translated and validated in Dutch. The VHI consists of 30 statements on voice-related aspects in daily life (with 5 response levels, scored 0 to 4). Summarising the scores on the 30 statements leads to a total VHI score, ranging from 0 to 120. A higher score corresponds to a worse voice-related functional status. Furthermore, the VHI includes an overall question on the quality of the voice with four response levels ranging from 0 (good), 1 (reasonable), 2 (moderate), 3 (poor). All VHI questionnaires were collected at baseline (i.e. before logopedic, surgical or medical voice treatment). To assess test–retest reliability, a subset of 30 patients (11 cancer, 13 structural lesion, 2 Reincke's oedema, 2 laryngitis, and 2 pareses) filled out the VHI twice, with a mean interval period of 3.5 months (range 1–6 months) without any voice intervention.

### Statistical analyses

Because of the skewed distribution of the VHI scores of the control group (the patient group showed normal distribution), independent Mann–Whitney tests ( $U$  test) and Kruskal–Wallis analysis-of-variance-by-ranks tests ( $H$  test) were used with a two-sided probability level of  $\leq 0.05$  to compare subject groups and to assess the association of VHI scores with age, gender, and self-reported voice quality.

The relations between VHI scores and case of voice impairment was evaluated with Receiver Operating Characteristics (ROC) analyses, using the area under the curve (AUC) as a summary measure of the overall discriminative ability of the VHI. In addition to ROC analyses, the sensitivity and specificity were calculated at various cut-off scores.

Internal consistency of the VHI was assessed by Cronbach's alpha. Test–retest stability was determined by Spearman's correlation coefficient between the first and the second (repeated) ratings. The clinical relevant difference score to be used in individual patients was defined as the maximum deterioration or improvement between test and retest scores. The clinically relevant difference score to be used in group study designs was defined based on an effect size (ES) of 0.80, being defined as the difference between the experimental group mean minus the control group mean divided by the standard deviation of the control group.

**Results**

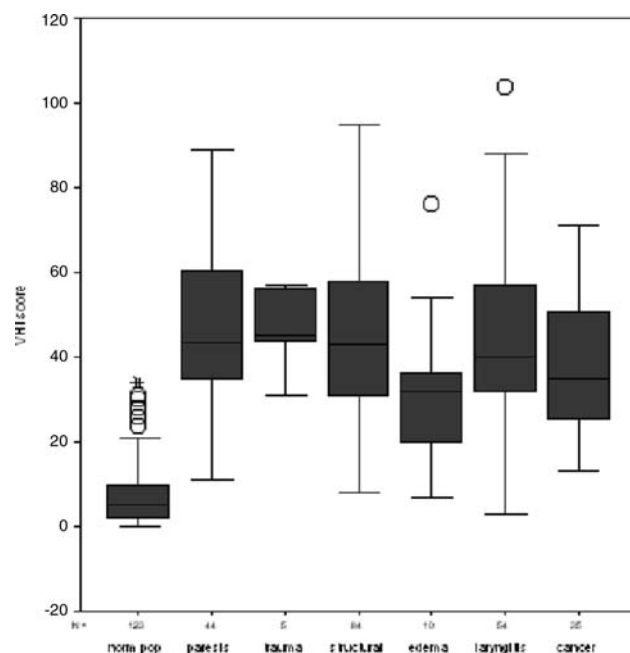
**Reliability**

Internal consistency of the VHI proved to be good with Cronbach’s alpha ranging from 0.87 (123 subjects from the normal population), 0.90 (35 glottic cancer patients), to 0.92 (196 voice-impaired patients), and 0.96 for the total group. Test–retest scores of the 30 patients who filled in the VHI twice over a mean period of 3.5 months (range 1–6 months) attested high test–retest stability with Spearman’s rho of 0.95 ( $P < 0.01$ ).

**Voice-impaired patients and the normal population**

Within the normal population 16% subjects judged their own voices as not good (score > 0 on the overall question on the quality of the voice) versus 93% of the patients with benign voice disorders and 94% of the cancer patients.

Voice handicap index scores of glottic cancer patients were similar to those of patients with voice problems due to benign lesions ( $P = 0.64$ ), but clearly deviant from the normal population ( $P < 0.01$ ) as were the scores of the total group of patients with benign voice disorders ( $P < 0.01$ ). An overview is given in Fig. 1. Because of this similarity between voice patient groups, further analyses were carried out on the total group of voice-impaired patients ( $n = 232$ ).



**Fig. 1** Boxplots presenting Voice handicap index scores for various subjects groups: normal population, patients with vocal fold paresis, larynx traumata, structural vocal fold lesions, Reincke’s oedema, laryngitis, and patients with voice problems after treatment for early glottic cancer

Sensitivity and specificity of the VHI in detecting voice-impaired patients using a range of cut-off points is shown in Table 1. The AUC was 0.98 (95% CI: 0.97–0.99) indicating good overall discriminative ability of the VHI. Table 2 shows that sensitivity and specificity is good with a cut-off point between 13 and 17. A cut-off point of 15 (or higher) on the VHI scale is proposed to identify patients with voice problems in daily life, because of a good degree of sensitivity and a sound (16% of the normal population judged their own voices as not-good) degree of specificity.

**Age, gender, and voice quality**

No association between the VHI scores with gender was found for the normal population ( $P = 0.86$ ) or for the voice-impaired patients ( $P = 0.59$ ).

Regarding age, no clear associations were present either in the normal population ( $r = 0.03$ ,  $P = 0.97$ ) or the voice-impaired patients ( $r = 0.01$ ,  $P = 0.99$ ).

Self-ratings of voice quality appeared to be clearly related to VHI scores with Spearman’s rho ranging from 0.32 for the normal population to 0.48 for the voice-impaired patients ( $P < 0.001$ ).

**Difference scores for individuals**

The difference score between the first and second rating appeared not to be dependent (Spearman’s  $r = -0.005$ ,  $P = 0.98$ ) on the height of the VHI score (Fig. 2). Individual difference scores between the first and second ratings remained within ten points, ranging from –9 to +10 points. Therefore a 10-point shift can be defined as a clinical relevant difference score to be used for single individual patients.

**Table 1** Overview of various Voice handicap index cut-off points regarding sensitivity and specificity

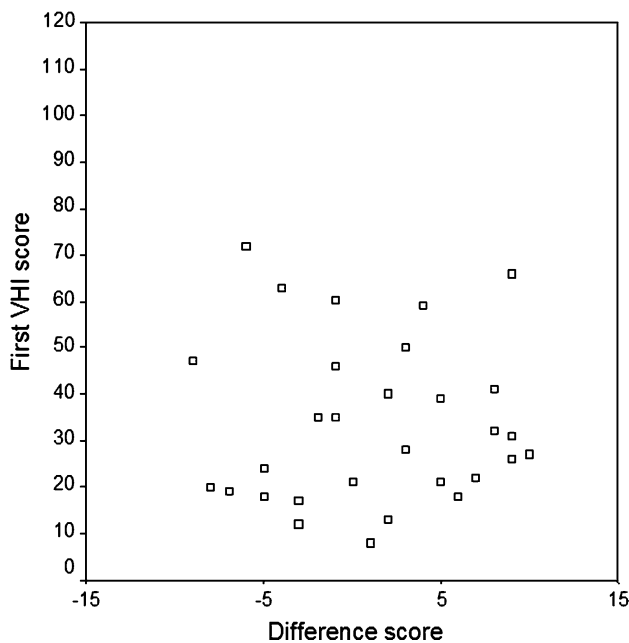
Cut-off point	Sensitivity (%)	Specificity (%)
12	98	79
13	97	82
14	97	84
<b>15</b>	<b>97</b>	<b>86</b>
16	96	88
17	95	92
18	94	92
19	92	93
20	89	94

In bold, the proposed cut-off point of 15 points or lower, which combines good Positive Predictive Value with high sensitivity and sound specificity

**Table 2** Overview of various Voice handicap index difference scores regarding effect size (with standard deviation of 19.40 as found in the total group of voice-impaired patients)

Difference score	Effect size
10	0.51
11	0.56
12	0.62
13	0.67
14	0.72
<b>15</b>	<b>0.77</b>
16	0.82
17	0.88
18	0.93

In bold, the proposed clinical relevant difference score of 15 points or more to be used in group-study designs



**Fig. 2** Scatter plot showing the (absent) relation between the first VHI score and the difference score between the first and second repeated VHI score as reported by 30 patients (Spearman's  $r = -0.005$ )

#### Difference scores for study designs

To define a relevant difference score for study designs with groups, determination of the effect size (ES) was used. ES above 0.80 represents a large statistical and clinical difference. From this study, standard deviations of the groups of voice-impaired patients at baseline ranged from 10.60 (trauma), 15.98 (early glottic cancer), 18.43 (paresis), 19.78 (structural lesions), to 20.50 (oedema and laryngitis); the standard deviation of the total group of voice-impaired patients ( $n = 232$ ) was 19.40. Table 2 represents an overview of effect sizes regarding various difference scores with

a standard deviation of 19.40 as representative for the total group of voice-impaired patients. The results show that a difference score of 15 points or more is clinically relevant in comparing groups of patients.

#### Discussion

The results of this study demonstrated a significant difference in mean VHI scores between patients with either benign voice pathology or voice pathology following treatment for glottic malignancy as compared to the normal population, which is in concordance with various previous studies. These studies, all about benign organic and/or functional voice disorders reported mean VHI scores varying from 11 to 47 which were found to differ significantly from controls with normal voices [4, 5, 11–13]. Nawka [11] was the first to report a significant difference between 9 patients with voice problems due to a malignant tumour (mean VHI score 34 points) and 16 normal control subjects (mean VHI score 7 points); moreover they also did not find a difference in VHI score between various diagnosis groups (benign organic or functional voice disorders ( $n = 159$ ), neurogenic voice disorders ( $n = 32$ ) or malignant voice disorders ( $n = 9$ )). From our results and the results as reported by Nawka et al., it is clear that voice problems in daily life of cancer patients are similar to those of patients with benign voice impairment. One could find this result remarkable because it might be expected that patients being cured of a malignancy experience the inherent voice impairment in a less negative way than patients cured of a benignancy.

The secondary aim of this study was to assess some underexposed psychometric characteristics of the VHI. Internal consistency proved to be good, as was test–retest stability. Regarding identification of voice-impaired patients, several authors used controls (subjects from the normal population without voice problems) in their randomised controlled studies on VHI change and reported mean “normal” values varying from 2.3 to 10.5 points but neither of them made a reliable effort to define a cut-off point [4, 5, 11, 12, 13]. The present study revealed a cut-off point of 15 to identify patients with voice problems in daily life.

Regarding clinical relevant difference scores, we found a difference score of 10 points to be useful for individuals in clinical practice and 15 points to be useful in study group designs. Jacobson et al. reported a shift of 18 points as a valuable difference score to measure efficacy of specific voice treatment techniques, but no clear analysis description was given [7]. Another non-statistical approach to define a clinically relevant difference score to be used in group design studies is to line up published studies on the efficacy of voice treatment and assess difference score appearing to be significant or non-significant. Four studies

on the efficacy of voice therapy in patients with several benign voice pathologies or voice pathologies following treatment for glottic malignancy, showed significant improvement of the mean VHI with a range of 12 to 18 points [10, 14, 15, 23]. On the contrary, Speyer [18] reported a non-significant median improvement of 6 points after voice therapy in patients with a diversity of chronic benign voice disorders. Other studies on the effect of several medical treatment modalities for different benign voice disorders show a mean VHI improvement ranging from 13 to 46 points [2, 13, 26]. All these studies on the efficacy of voice intervention on various voice patient groups reveal that a statistical difference score is at least 12 points. A meta-analysis could provide further information but it seems too early to perform such a study because of the limited number of studies on efficacy of voice treatment at this moment. In the mean time, we propose a difference score of 15 points signifying a statistical and clinical high effect size.

The proposed cut-off point and the clinical difference scores in this study are not meant to be conclusive, mainly because of the Dutch origin of the data, which may have influenced the results. Currently, a European VHI Study Group is working on comparison of various translations of the VHI to assess equivalence. The first preliminary results reveal that there are only minor differences between the included versions, but further data exploring is ongoing.

## Conclusion

Patients with voice problems after treatment for early glottic cancer encounter the same amount of problems in daily life as other voice-impaired patients and therefore require the same attention and care for this sequel to their initial cancer treatment. Furthermore, the VHI proved to be an adequate tool for baseline and effectiveness measurement of voice.

## References

- Aref A, Dworkin J, Devi S, Denton L, Fontanesi J (1997) Objective evaluation of the quality of voice following radiation therapy for T1 glottic cancer. *Radiother Oncol* 45(2):149–153
- Benninger MS, Gardner G, Grywalski C (2001) Outcomes of botulinum toxin treatment for patients with spasmodic dysphonia. *Arch Otolaryngol Head Neck Surg* 127:1083–1085
- Fung K, Yoo J, Leeper HA, Hawkins S, Heeneman H, Doyle PC, Venkatesan VM (2001) Vocal function following radiation for non-laryngeal versus laryngeal tumors of the head and neck. *Laryngoscope* 111:1920–1924
- Guimares I, Abberton E (2004) An investigation of the voice handicap index with speakers of Portuguese: preliminary data. *J Voice* 18(1):71–82
- Günther S, Rasch T, Klotz M, Hoppe U, Eysholdt U, Rosanowski F (2005) Voice related quality of life and voice handicap index in dysphonic patients. A comparison (German) *HNO* 53(10):895–904
- Hocevar -Boltežar I, Žargi M (2000) Voice quality after radiation therapy for early glottic cancer. *Arch Otolaryngol Head Neck Surg* 126:1097–1100
- Jacobson BH, Johnson A, Grywalski C, Sibergleit A, Jacobson G, Benninger MS, Newman CW (1997) The voice handicap index (VHI): development and validation. *Am J Speech Lang Pathol* 6:66–70
- Jepsen MC, Gurushanthaiah D, Roy N, Smith ME, Gray SD, Davis RK (2003) Voice, speech, and swallowing outcomes in laser-treated laryngeal cancer. *Laryngoscope* 113:923–928
- Lehman JJ, Bless DM, Brandenburg JH (1988) An objective assessment of voice production after radiation therapy for stage I squamous cell carcinoma of the glottis. *Head Neck Surg* 98:121–129
- Machulla R, Hacki T, Hoppe U, Rosanowski F (2005) Voice Handicap Index in rehabilitation. *HNO* Mar 2, [Epub ahead of print] German
- Nawka T, Wiesmann U, Gonnermann U (2003) Validierung des voice handicap index in der deutschen fassung. *HNO* 51:921–929
- Peeters AJGE, Van Gogh CDL, Goor KM, Verdonck-de Leeuw IM, Langendijk JA, Mahieu HF (2004) Health status and voice outcome after treatment for T1a glottic carcinoma. *Eur Arch Otorhinolaryngol* 261:534–540
- Rosen CA, Lee AS, Osborne J, Zullo T, Murry T (2004) Development and validation of the voice handicap index-10. *Laryngoscope* 114:1549–1556
- Rosen CA, Murry T, Zinn A, Zullo T, Sonbolian M (2000) Voice handicap index change following treatment of voice disorders. *J Voice* 14:619–623
- Roy N, Gray SD, Simon M, Dove H, Corbin-Lewis K, Stemple JC (2001) An evaluation of the effects of two treatment approaches for teachers with voice disorders: a prospective randomized clinical trial. *J Speech Lang Hear Res* 44:286–296
- Rydell R, Schalen L, Fex S, Elnér Å (1995) Voice evaluation before and after laser excision vs. radiotherapy of T1a glottic carcinoma. *Acta Otolaryngol (Stockh)* 115:560–565
- Smith JC, Johnson JT, Cognetti DM, Landsittel DP, Gooding WE, Cano ER, Myers EN (2003) Quality of life, functional outcome, and costs of early glottic cancer. *Laryngoscope* 113:68–76
- Speyer R, Wieneke GH, Dejonckere PH (2004) Self-assessment of voice therapy for chronic dysphonia. *Clin Otolaryngol* 29(1):66–74
- Stewart MG, Chen AY, Stach CB (1998) Outcomes analysis of voice and quality of life in patients with laryngeal cancer. *Arch Otolaryngol Head Neck Surg* 124:143–148
- Stoeckli SJ, Guidicelli M, Schneider A, Huber A, Schmid S (2001) Quality of life after treatment for early laryngeal carcinoma. *Eur Arch Otorhinolaryngol* 258:96–99
- Van der Torn M, Verdonck-de Leeuw IM, Kuik DJ, Mahieu HF (2002) Communicative suitability of voice following radiotherapy for T1a glottic carcinoma: testing the reliability of a rating instrument. *J Voice* 16:398–407
- Van Gogh CDL, Verdonck-de Leeuw IM, Boon AB, Langendijk JA, Kuik DJ, Mahieu HF (2005) A screening questionnaire for voice problems after treatment for early glottic cancer. *Int J Radiat Oncol Biol Phys* 62:700–705
- Van Gogh C, Verdonck-de Leeuw I, Boon-Kamma A, Rinkel RN, de Bruin MD, Langendijk JA, Kuik DJ, Mahieu HF (2006) Efficacy of voice therapy in patients after treatment of early glottic cancer. *Cancer* 106(1):95–105
- Verdonck-de Leeuw IM, Hilgers FJM, Keus RB, Koopmans-van Beinum FJ, Greven AJ, de Jong JM, Vreeburg G, Bartelink H (1999) Multidimensional assessment of voice characteristics after radiotherapy for early glottic cancer. *Laryngoscope* 109:241–248

25. Verdonck-de Leeuw IM, Keus RB, Hilgers FJM, Koopmans-van Beinum FJ, Greven AJ, de Jong JM, Vreeburg G, Bartelink H (1999) Consequences of voice impairment in daily life for patients following radiotherapy for early glottic cancer: voice quality, vocal function, and vocal performance. *Int J Radiat Oncol Biol Phys* 44:1071–1078
26. Wingate JM, Ruddy B, Lundy DS, Lehman J, Casiano R, Collins SP, Woodson GE, Sapienza C (2005) Voice handicap index results for older patients with adductor spasmodic dysphonia. *J Voice* 19(1):124–131