Voice Handicap Index and Interpretation of the Cutoff Points Using Receiver Operating Characteristic Curve as Screening for Young Adult Female Smokers

***Dionysios Tafiadis, †Evangelia I. Kosma, ‡Spyridon K. Chronopoulos, §Aggelos Papadopoulos, *Konstantinos Drosos, *Vassiliki Siafaka, *Eugenia I. Toki, and *Nausica Ziavra,** **Ioannina, †Larissa, ‡Arta, and §Patra, Greece*

Summary: The relationship between smoking and alterations of the vocal tract and larynx is well known. This pathology leads to the degradation of voice performance in daily living. Multiple assessment methods of vocal tract and larynx have been developed, and in recent years they were enriched with self-reported questionnaires such as Voice Handicap Index (VHI). This study determined the cutoff points of VHI's total score and its three domains for young female smokers in Greece. These estimated cutoff points could be used by voice specialists as an indicator for further clinical evaluation (foreseeing a potential risk of developing a vocal symptom because of smoking habits). A sample of 120 female nondysphonic smokers (aged 18–31) was recruited. Participants filled out the VHI and Voice Evaluation Form. VHI's cutoff point of total score was calculated at the value of 19.50 (sensitivity: 0.780, specificity: 0.133). Specifically, the construct domain of functional was 7.50 (sensitivity: 0.900, specificity: 0.217), for physical it was 8.50 (sensitivity: 0.867, specificity: 0.483), and for emotional it was 7.50 (sensitivity: 0.833, specificity: 0.200) through the use of receiver operating characteristic. Furthermore, VHI could be used as a monitoring tool for smokers and as a feedback for smoking cessation.

Key Words: Voice Handicap Index–Females–Cutoff scores–Smoking cessation–Screening.

INTRODUCTION

Voice disorders can emerge from alterations of the voice mechanism because of smoking.[1–3] The relationship between smoking and changes in the vocal tract and larynx has been well documented.[4–6] This documentation includes lesions such as vocal polyps,[4] Reinke's edema,[7] vocal fold carcinomas,[6,8,9] chronic inflammation, erythema or irritation of laryngeal mucosa,[9–11] leading to changes of voice's acoustic characteristics.[12–15] Many people (including smokers) seek consultation and voice evaluation as soon as they experience any type of vocal symptom.

Evidence-based models[16,17] and protocols of voice evaluation include laryngeal imaging,[18–22] and acoustic-perceptual[23–26] and aerodynamic assessments.[27–29] However, in recent years these protocols have been enriched with self-reported questionnaires[30] that determine the patients' perception of their voice status. The outcomes of questionnaires specify the degree of the disorder's perceived severity as well as the impact of voice disorders on the quality of life.[23–26] Consequently, standardized self-report questionnaires have been included in the overall screening procedure.[31,32] The most frequently used tool for research purposes is the Voice Handicap Index (VHI)[33] that has been translated in many languages.[34–43]

VHI is a self-reporting tool consisting of three different domains for evaluating the functional, physical, and emotional aspects of voice. Each domain includes 10 items (questions) that correspond to a total number of 30. Every item ranges from 0 = never to 4 = always, on a Likert-type scale. After the development of the original VHI, its different types and forms have been released for both pediatric[44] and adult populations.[45–47] Furthermore, VHI has been applied in different types of dysphonic populations[48,49] such as adductor's spasmodic dysphonia[50,51] and different vocal fold lesions.[52,53] Furthermore, it has been used to evaluate professional voice users[54–60] and smokers.[61,62]

Moradi and colleagues suggested a cutoff point of VHI for voiced disordered Persian population using receiver operating characteristic (ROC).[62] The ROC methodology can be dated back to the 1950s as it primarily evolved from radar signal's evaluations and for tracking electronic signals. Particularly, the methodology produces positive or negative results in conjunction with the presence or absence of a disease. The overall accuracy of the procedure is described by "sensitivity" and "specificity," which are the probabilities of a result being positive (presence of the disease) or negative (absence of the disease), respectively. In general, sensitivity reflects the amount of patients positively diagnosed with a disease whereas specificity measures the false positive rate. Moreover, a cutoff point is determined by the closest point near the upper left corner of a positive curve and vice versa. According to aforementioned, the resulting curve (ROC) presents sensitivity in relevance to specificity while giving more clues toward a well-estimated prediction through paired information (sensitivity-specificity).[63] This type of curve is vastly used in medical decision-making, as well as in data mining and machine learning processes.[64] If this curve passes near the upper corner, then both sensitivity and specificity are equal to 100%, whereas the curve's performance deteriorates at a diagonal line of 45° (y = x), which will act as a random guessing operation. The points of (0, 0) and (1, 1) show two different key points that are named "no true positives with no errors" and "100% true positives with vast errors," respectively.

Accepted for publication March 13, 2017.
Conflict of interest: All authors declare no conflicts of interest in this manuscript.
From the *Department of Speech and Language Therapy, Technological Educational Institute of Epirus, Ioannina, Greece; †Faculty of Medicine, School of Health Sciences, University of Thessaly, Biopolis, Larissa, Greece; ‡Department of Computer Engineering, Technological Educational Institute of Epirus, Arta, Greece; and the §Department of Paediatrics, Karamandanio General Children's Hospital NHS, Patra, Greece.
Address correspondence and reprint requests to Dionysios Tafiadis, Department of Speech and Language Therapy, Technological Educational Institute of Epirus, 4th Km Ioaninna—Athens Road, Ioannina, P.C. GR455 00, Greece. E-mail: d.tafiadis@ioa.teiep.gr; tafiadis@gmail.com
Journal of Voice, Vol. ■■, No. ■■, pp. ■■-■■
0892-1997

http://dx.doi.org/10.1016/j.jvoice.2017.03.009

In other words, the diagonal line includes points corresponding to random guesses. If the type of line includes points of (0,0) and points near (0,1), then it is considered as a perfect classification. Also, an additionally important factor is the calculation of the area under the curve (AUC), which is always estimated to be between a value of 0 and 1.[65] This area can be split into four categories with predictive abilities named as "No predictive" (AUC < 0.5), "Acceptable" (0.7 to less than 0.8), "Excellent" (0.8 to less than 0.9), and "Outstanding" (≥0.9).[66] Also, Wilcoxon rank sum tests[65] are equivalent to AUC[67,68] and have almost the same value as Gini coefficient[67] that corresponds to the twice area between the diagonal and the ROC curve. Also, unity equals to the result of 2AUC − Gini.[68]

The purpose of the current study was to calculate the four cutoff points of VHI's total score and its three domains. These cutoff points will serve as indicators for further referral to voice clinicians.

MATERIALS AND METHODS

Participants

One hundred twenty young female speech therapy students (60 smokers and 60 nonsmokers) were recruited for this study. All participants did not experience any laryngeal or respiratory disorders in the last 2 weeks before enrollment. Also, they did not have any history of alcohol or drug addiction, consistent symptoms of gastroesophageal reflux disease or laryngopharyngeal reflux, voice misuse, and reported environmental factor that could deteriorate their voices' condition.

Data collection

All students filled out the Voice Evaluation Form (VEF)[69] and Hellenic VHI.[42] VEF is a consensus template (including over 70 items/questions) that was developed by the American Speech Hearing Association to copy the current and the former background of an examinee's voice status. VEF served as a voice history form for this study.

The Greek standardized version of VHI[43] was also administrated. VHI consisted of 30 questions, which were summarized to a total score (VHI-T) split into three domains: emotional (VHI-E), physical (VHI-P), and functional (VHI-F). Each domain included 10 questions (30 items in total).

The research was approved by the Department of Speech and Language Therapy Research Ethic Committee. The data collected in the School of Health and Welfare of the Technological Educational Institute of Epirus. All participants were informed of the research purposes and signed a consensus letter. They were also informed that their personal data will remain confidential. After consenting to the research, all participants were asked to fill in the two questionnaires.

Statistical analysis

Variables with skewed distribution were expressed as median (interquartile range). Mann-Whitney *U* test was used for the comparison of continuous variables between the two study groups. Also, the ROC curve was used for the assessment of the cutoff value for VHI. All reported *P* values were two-tailed. Statistical significance was set at $P < 0.05$ and analysis was conducted using SPSS statistical software (version 19.0, Armonk, NY, USA).

RESULTS

Sample consisted of 120 female speech therapy students (60 smokers and 60 nonsmokers) as already mentioned. Total samples' mean age was 22.32 years (standard deviation [SD] = 2.49) ranging from 18 to 34. The average years of smoking were 2.23 (SD = ±1.35) and the number of cigarettes per day was equal to a mean value of 15.21 (SD = 8.23).

Smokers had a significant higher overall VHI median (median = 29.00) compared to nonsmokers (median = 16.00), $U = 358.00$, $P < 0.001$. Same statistically significant differences of medians were computed for VHI-F (smokers median = 9.50 and nonsmokers median = 5.00, $U = 469.00$, $P < 0.001$), VHI-P (smokers median = 9.00 and nonsmokers median = 6.00, $U = 943.00$, $P < 0.001$), and VHI-E (smokers median = 10.00 and nonsmokers median = 6.00, $U = 426.00$, $P < 0.001$). The smokers exhibited the higher scores (Table 1).

A ROC analysis was conducted to determine whether the three domains of VHI-T and VHI could estimate the starting point of possible pathology due to smoking. The analysis revealed statistically significant positive discrimination between smokers and nonsmokers. A high effect value was noted for VHI-T (AUC 0.899, $P < 0.001$), whereas good effects were noted for VHI-F (AUC 0.868, $P < 0.001$) and VHI-E (AUC 0.880, $P < 0.001$). Mediocre effect was noted for VHI-P domain (AUC 0.734, $P < 0.001$) (Table 2).

ROC curve test determined the cutoff points for VHI-T score and the three domains of VHI. The cutoff score for VHI-T was positive if greater than or equal to 19.50 with sensitivity of 0.780 and specificity of 0.133 (Figure 1). VHI-F cutoff point was

TABLE 1.
Comparison of Medians Between Female Smokers and Nonsmokers for VHI Total Score and VHI Domains

	Smokers (N = 60)	Nonsmokers (N = 60)		
	Median (IQR)	Median (IQR)	Mann-Whitney *U*	*P* Level
Total	29.00 (21.50–35.00)	16.00 (14.00–19.00)	358.00	<0.001*
Functional	9.50 (8.00–12.00)	5.00 (3.00–6.00)	469.00	<0.001*
Physical	9.00 (6.00–12.00)	6.00 (4.00–7.00)	943.00	<0.001*
Emotional	10.00 (8.00–13.00)	6.00 (5.00–7.00)	426.00	<0.001*

* *P* level at <0.05.
Abbreviations: IQR, interquartile range; VHI, Voice Handicap Index.