

# Voice Acoustic Analysis of Pediatric Vocal Nodule Patients Using Ratios Calculated With Biomedical Image Segmentation

\*Nagihan Bilal, †Turab Selcuk, \*Selman Sarica, †Ahmet Alkan, \*İsrafil Orhan, ‡Adem Doganer, \*Saime Sagioglu, and §Mehmet Akif Kilic, \*†‡Kahramanmaraş and §Istanbul, Turkey

**Summary: Objective.** The aim of this study was to determine nodules using newly developed software with a computer-assisted visual process technique for the calculation of size. The effects of the ratios of nodule base and width were evaluated with voice acoustic analysis.

**Methods.** A total of 72 patients with pediatric vocal nodule were evaluated. Nodules were marked with the *ImageJ News* program on photographs obtained from the video recordings in the videostroboscopic examination and classified according to the Shah et al scale. Segmentation was applied automatically. The ratios were taken as base of nodule/width and base of nodule/vocal cord. In the voice acoustic analysis, basic frequencies (mean F0), jitter (local %), shimmer (local %), and harmonicity (mean harmonics-to-noise [mean HNR]) were evaluated.

**Results.** A statistically significant negative correlation was determined between the mean F0 value and the nodule base/width ratio ( $P = 0.042$ ,  $r = -0.240$ ). A negative statistically significant relationship was determined between jitter (%) and vocal nodule base/width ( $P = 0.009$ ,  $r = -0.305$ ). A statistically significant positive correlation was determined between mean HNR and vocal nodule base/width ( $P = 0.034$ ,  $r = 0.324$ ). In discriminant analysis, correct classification of the Shah et al scale degrees of the classifying variables was 73.6%.

**Conclusion.** Through collaboration with the biomedical engineering department, the results of this study determined new ratios in patients with pediatric vocal nodule. In voice acoustic analysis, the mean F0 was more affected by the width of the nodule, mean HNR was affected by the length of the base of the nodule, and jitter (%) was affected by the width of the nodule.

**Key Words:** Child–Vocal fold nodule–Voice quality–Laryngeal image analyses–Base of the nodule.

## INTRODUCTION

Pediatric vocal nodule (PVN) is the disease that most commonly caused dysphonia in childhood. In 40% of cases with voice disorders, a diagnosis of nodule is made.<sup>1</sup> Laryngostroboscopy is the best method for diagnosis and for the evaluation of morphodynamic changes. Nodules are seen as submucosal lesions at the junction of the anterior and midsection of the vocal cords.<sup>2</sup> PVN develops as a result of trauma, which occurs because of overuse of the voice, repeated vocal abuse, and vocal strain. Contact between the two opposing surfaces of the vocal cords causes trauma.<sup>3</sup> Acoustic measurements are another evaluation

method in patients with PVN. The diagnosis of PVN is problematic as there are more subjective evaluation methods than objective grading systems.<sup>4</sup>

As the grading systems in the literature are limited by the clinician's experience and knowledge, a completely objective evaluation cannot be made.<sup>1,5</sup> These grading systems that have been used in the literature are related to the measurement of voice quality with PVN dimensions<sup>3,6</sup> and it is noticeable that there has been no new evaluation method developed for the measurement of voice quality with vocal nodule size. As a subjective system is dependent on the physician's knowledge and experience, the existence of an objective system is important in the follow-up of treatment and to be able to objectively show the success of the treatment methods used. There is a need for this system for the more objective evaluation of healing after voice therapies.

In previous studies, PVNs have always been evaluated with subjective measurements. In this study, nodules were first identified with software that determined PVNs. The primary aim of this study was to create a computer program using computer-assisted imaging techniques, to determine nodules, calculate nodule size, and determine any changes in the voice analysis related to these dimensions.

## MATERIALS AND METHODS

Approval for the study was granted by the Ethics Committee of Sıtcu Imam University of Medical Sciences (No. 225, dated: August 25, 2016). Informed consent was obtained from the parent or legal guardian of each child included in the study.

Accepted for publication November 16, 2017.

Author contributions: Nagihan Bilal: Conception, data acquisition, data interpretation, drafting, revising content, final approval, accountability for all aspects of work. Turab Selcuk: Conception, data interpretation, drafting, revising content, final approval, accountability for all aspects of work. Selman Sarica: Data acquisition, drafting, final approval, accountability for all aspects of work. Ahmet Alkan: Conception, data interpretation, revising content, final approval, accountability for all aspects of work. İsrafil Orhan: Conception, revising content, final approval, accountability for all aspects of work. Adem Doganer: Evaluation of statistical analyses. Saime Sagioglu: Conception, data acquisition, data interpretation. Mehmet Akif Kilic: Conception, data acquisition, data interpretation, drafting, revising content, final approval, accountability for all aspects of work.

From the \*Department of Otorhinolaryngology, Faculty of Medicine, Kahramanmaraş Sıtcu Imam University, Kahramanmaraş, Turkey; †Faculty of Engineering and Architecture Electric Electronic Engineering, Kahramanmaraş Sıtcu Imam University, Kahramanmaraş, Turkey; ‡Department of Biostatistics and Medical Informatics, Faculty of Medicine, Kahramanmaraş Sıtcu Imam University, Kahramanmaraş, Turkey; and the §Department of Otorhinolaryngology, Faculty of Medicine, Medeniyet University, İstanbul, Turkey.

Address correspondence and reprint requests to Nagihan Bilal, Department of Otorhinolaryngology, Faculty of Medicine, Kahramanmaraş Sıtcu Imam University, HaydarBeyMah 32120 Sok Beytepesitesi4 Blok No:13, Kahramanmaraş, Turkey. E-mail: nagihanyazan@gmail.com

Journal of Voice, Vol. 33, No. 2, pp. 195–203

0892-1997

© 2017 The Voice Foundation. Published by Elsevier Inc. All rights reserved.

<https://doi.org/10.1016/j.jvoice.2017.11.010>

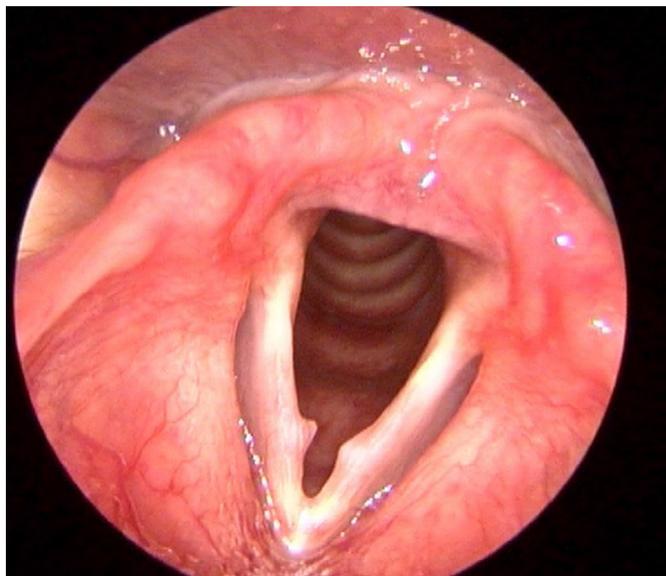
### Study participants

The study included patients aged 0–18 years who presented at the phoniatory polyclinic and were diagnosed with vocal nodules between August 2016 and March 2017. Voice recordings were made in a silent room under the guidance of a speech language pathologist. A Shure SM58 dynamic microphone (Shure Inc, Niles, IL) was used for the recording. During the voice recording with *Adobe Audition 3.0* software (Adobe Systems Inc., San Jose, CA, USA), the microphone was held 5 cm from the mouth in a vertical position. Patients were audio-recorded while producing the vowel /â/ at a comfortable volume and a constant pitch for as long as possible. This was repeated twice with about 5 seconds between trials. The settings for the Adobe Audition were mono, 16-bit resolution, and sampling speed of 44,100 Hz. A slice of approximately 2 seconds was taken from the recorded voice for acoustic analysis. An interval of 2 seconds from the midportion of each sample was selected for acoustic analysis. The obtained slices were recorded in PCM wave format. Videostroboscopic (fiberoptic nasolaryngoscope; Olympus, Tokyo, Japan) recordings were taken.

Patients were excluded if it was not possible to take jitter, shimmer, and mean F0 measurements; if they had an upper or lower respiratory tract infection; any systemic disease that could affect voice quality; and any history of upper respiratory tract surgery or voice therapy. Furthermore, male patients in puberty (14–17 years old) were not included in the study. Those patients with incomplete examination records, insufficient voice recordings or videostroboscopic examination records, or unclear photographic images were also excluded from the study.

A total of 145 patients were initially examined, and after the exclusion of 73 patients with unsuitable voice recordings, the study evaluation was completed with a total of 72 patients. All patients included in the study were evaluated by one ear, nose, and throat physician and one speech language pathologist. The patient records were evaluated by two ear, nose, and throat physicians. The video recordings of the patients were examined and photographs were taken from the recordings (HP computer Z400 6-DIMM Workstation; Hewlett-Packard, Japan). To prevent the image of false vocal cords, the images included in the study were those from above and fully centered. The drawing of the images was decided in the *ImageJ* program (Bethesda, Maryland, USA). Several photographs were taken from the video images. The images selected were those showing the bilateral vocal nodules most clearly. Photographs from the recordings of the vocal cords open (completely abducted) and closed were obtained with the *Pinnacle Studio Ultimate Collection* version 14 program (Pinnacle Systems Ltd, Iver Heath, UK) (*Image 1*).

From the photographs and video recordings obtained, each patient was classified according to the Shah et al scale, and the grades were recorded.<sup>1</sup> The measurements were planned taking the Shah et al. scale as reference for the measurements of PVN dimensions, and the grades were determined for each nodule according to this scale.<sup>1</sup> When determining the size of the nodules, discrete nodules were labeled as “d” and sessile nodules as “s.” According to the scale, those with a wide base and anterior posterior extension, in other words, those with the base at least twice



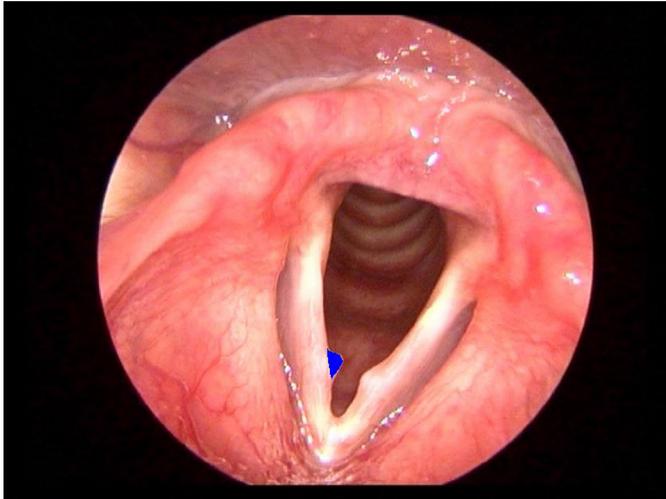
**IMAGE 1.** Photograph obtained from the video records after examination of the patient.

the width (extension from lateral to medial) were named sessile nodules.

The nodules on these recordings were marked with the *ImageJ* News (1.51k) program and these images were recorded. In the program, the vocal nodule was identified first by drawing around it on the images (*Image 2*). Then the internal section was filled (*Image 3*) and areas outside the vocal nodule were erased by the program (*Image 4*). In the second stage, the vocal cord on the side where the nodule was located was marked, the inside was filled, and the surrounding areas were erased (*Image 5*). These images were recorded. The drawings were made separately on the right and left side nodules. On these created images, the base of the nodule (*Image 6*) and the width (*Image 7*) were measured and calculated as pixels<sup>2</sup>. Ratios were calculated as base/width and base/vocal cord. The average of the bilateral nodules was taken. In the same way, the sections remaining between the vocal cords when they were completely closed were drawn in



**IMAGE 2.** Drawing around the nodule in the ImageJ program.



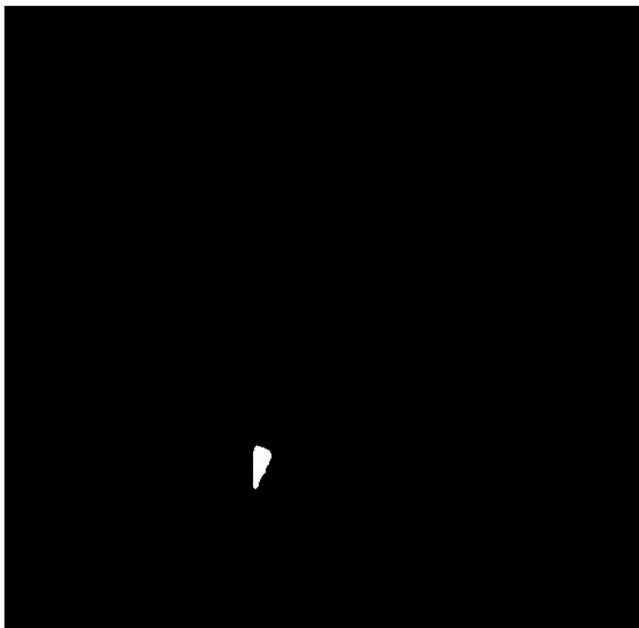
**IMAGE 3.** The inside of the nodule in the ImageJ program.

the *ImageJ News* program. When closed, the gap between the glottis was marked and the base/width ratio was taken from here.

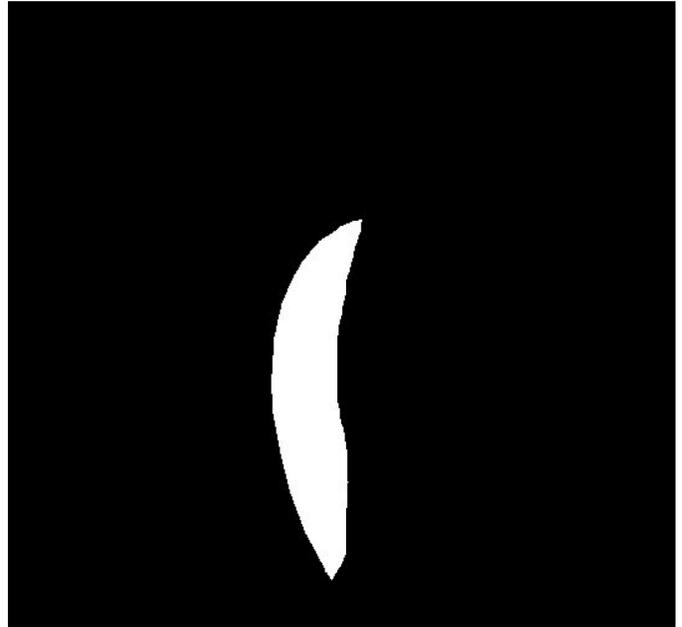
Voice analysis was made with the *Praat* program (version 5.3.53, Boersma & Weenink 2013).<sup>7</sup> The values for the basic frequencies (mean F0), jitter (local %), shimmer (local %), and harmonicity (mean harmonics-to-noise [mean HNR]) were taken.

#### Automatic segmentation of nodules

In the study, the vocal cord images of 72 patients were used as the data set. The image resolution was  $580 \times 780$  pixels<sup>2</sup>. Before the image processing phase, a rectangular region was defined that could include the vocal cord in all images in the data set. At this point, the image processing algorithms were applied only to this region. In the image preprocessing phase, first, each of the images was converted to a gray level image and then the

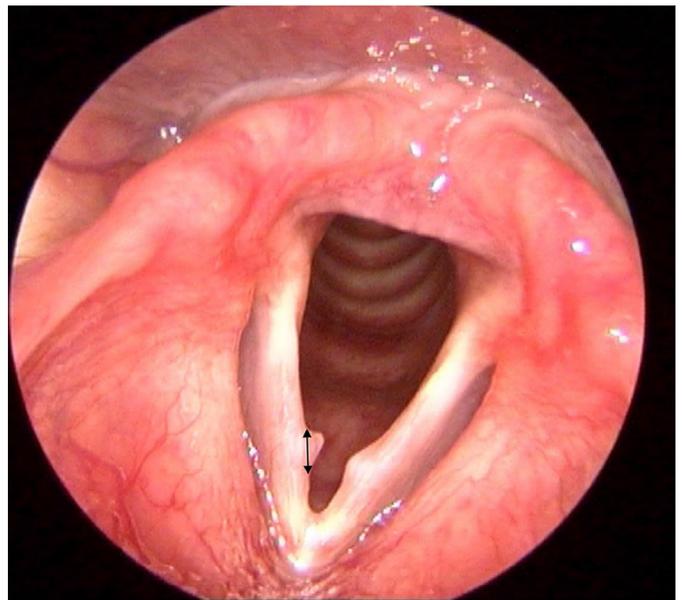


**IMAGE 4.** The appearance of the nodule after erasure of the surrounding tissues in the ImageJ program.

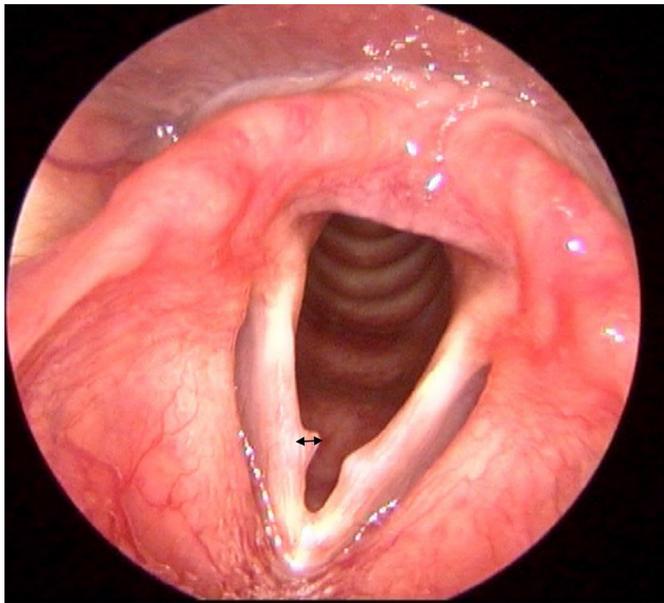


**IMAGE 5.** Drawing of the vocal cord with the ImageJ program.

contrast was enhanced by preprocessing with the adaptive histogram equalization method (*Matlab Version 9.2* [R2017a, Natick, MA, USA]). A histogram is a graph showing the gray value of the pixels in the image. Adaptive histogram equalization improves the visibility of the glottis. In the processing phase, the first derivative of the image was taken and edge regions with sudden gray change were obtained. Then the image was converted to a black and white image with the OTSU thresholding method.<sup>8</sup> As there was the same edge information of the glottis and the vocal cord, glottosis segmentation was performed in the study. For this, the connected component method, which detects the largest white pixel group in the image, was used.



**IMAGE 6.** Arrow indicating the base of the vocal nodule.



**IMAGE 7.** Width of the vocal nodule.

Thus, both the boundary of the glottis and the boundary of the nodules were obtained and nodules were segmented. The images of the processing steps are shown in [Figure 1](#).

### Statistical analyses

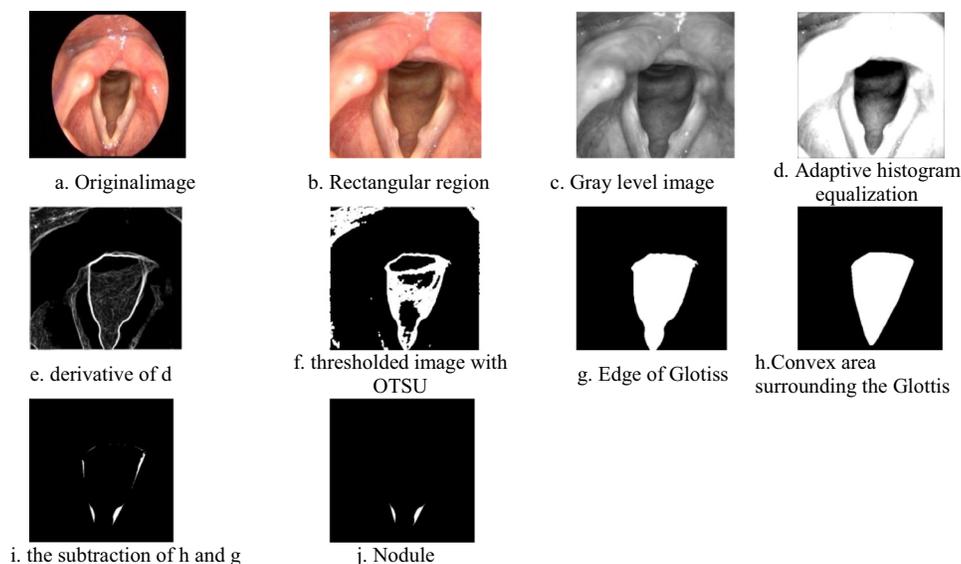
Statistical analyses of the study data were made with *SPSS* version 22 software (IBM SPSS for Windows, Version 22, IBM Corporation, Armonk, NY). Conformity of the variables to normal distribution was examined with the Kolmogorov-Smirnov test. In the comparison of multiple groups, one-way variance

analysis was used, and for post-hoc comparisons, the Tukey test was used. In the comparison of groups not showing normal distribution, the Kruskal-Wallis H test was applied. The relationship between variables was examined with the Spearman correlation test. The Shah et al scale classification, classified according to the variables, was examined with discriminant analysis. Statistical parameters were stated as mean  $\pm$  standard deviation (SD). A value of  $P < 0.05$  was accepted as statistically significant.

### RESULTS

The vocal nodules of a total of 72 patients, comprising 22 (30.6%) female and 50 (69.4%) male, with a mean age of  $11.33 \pm 2.27$  years, was evaluated. Of the male patients, 52% were in the 10–13 years age group and 40.9% of the female patients were in the 14–17 years age group. In all the patients, the vocal cord nodules were observed to be bilateral and in the area of the junction of the anterior third and the posterior two-thirds. The demographic details of the patients are given in [Table 1](#). According to the Shah et al scale, 19 patients (26.4%) were classified as grade 1s, 25 (34.7%) as grade 2s, 14 (19.4%) as grade 3s, 1 (1.4%) as grade 1d, 9 (12.5%) as grade 2d, and 4 (5.6%) as grade 3d.

In the results of the voice analysis of the patients, the mean F0 was  $245.22 \pm 53.38$  (min–max: 102.64–358.20), jitter (local %)  $0.75 \pm 0.84$  (min–max: 0.20–6.56), shimmer (local %)  $5.41 \pm 4.25$  (min–max: 1.44–32.38), and mean HNR  $16.95 \pm 4.85$  (min–max: 0.98–26.70). The results of the voice analyses of the patients are shown in [Table 2](#). No statistically significant difference was determined between the genders with respect to the findings of the voice analysis. The results of the voice analyses of the patients according to the Shah et al. scale are shown in [Table 3](#). In the evaluation of the voice analyses of patients with



**FIGURE 1.** The stages from the imaging process to the segmentation of the nodules. First, the color image shown in **b** is converted to the gray level image (**c**). In the gray level image, the contrast is improved by adaptive histogram equalization method (**d**) and the first derivative of the image is taken (**e**). Thus, the edges became obvious. It was then thresholded by the OTSU method (**f**). And then the inside of the glottis was filled (**g**) and component labeling method was applied. Thus, a polygon surrounding the glottis was obtained (**h**). When this polygon is differentiated from glottis, nodules and similar regions are obtained (**i**). Finally, nodule segmentation was performed using geometric features of the nodule (**j**).

**TABLE 1.**  
**Characteristics of the Patients According to Age and Gender**

	N	Mean Age (y)	SD	Min	Max
Total	72	11.33	2.27	5.00	15.00
Male	50 (69.4%)	11	1.91	5.00	13.00
Female	22 (30.6%)	12.09	2.83	6.00	15.00

PVN according to the Shah et al scale, a statistically significant difference was determined between all the groups with respect to mean F0 ( $P = 0.015$ ).

The highest mean F0 value was determined at 3d. No statistically significant difference was determined between the genders according to the Shah et al scale ( $P = 0.271$ ).

The base/width ratio of the vocal cord nodules was determined as mean  $3.86 \pm 0.87$  (min-max: 1.83–5.87) and the base/vocal cord ratio as  $0.21 \pm 0.05$  (min-max: 0.08–0.37) (Table 2). No statistically significant difference was determined between the genders with respect to these ratios. A statistically significant difference was found between the Shah et al scale values

**TABLE 2.**  
**Mean F0, Jitter, Shimmer, and Mean HNR Values**

	Gender				<i>P</i>
	Male		Female		
	Mean $\pm$ SD	(Min–Max)	Mean $\pm$ SD	(Min–Max)	
Mean F0	235.76 $\pm$ 44.29	(102.64–289.00)	255.69 $\pm$ 55.89	(132.68–358.20)	0.484
Jitter	0.57 $\pm$ 0.25	(0.20–1.28)	0.94 $\pm$ 1.33	(0.27–6.56)	0.115
Shimmer	4.69 $\pm$ 2.47	(1.44–11.08)	6.13 $\pm$ 6.86	(1.50–32.38)	0.187
Mean HNR	16.81 $\pm$ 4.01	(9.44–26.60)	17.10 $\pm$ 6.62	(0.98–26.70)	0.525
Base/Width	3.82 $\pm$ 0.80	(2.38–5.87)	3.95 $\pm$ 1.04	(1.83–5.28)	0.568
Base/Vocal fold	0.21 $\pm$ 0.05	(0.13–0.37)	.20 $\pm$ 0.05	(0.08–0.31)	0.795
Posterior glottic gap ratio	1.17 $\pm$ 2.46	(0.00–11.63)	1.79 $\pm$ 2.55	(0.00–7.25)	0.330

Notes: Independent samples *t* test;  $\alpha$ : 0.05.

**TABLE 3.**  
**Comparison of the Measurement Values According to the Shah et al Scale**

	Scale From Shah et al						<i>P</i>
	1s	2s	3s	1d	2d	3d	
	Mean $\pm$ SD	Mean $\pm$ SD	Mean $\pm$ SD	Mean	Mean $\pm$ SD	Mean $\pm$ SD	
<b>Mean F0*</b>	254.22 $\pm$ 55.43	219.29 $\pm$ 52.59	274.14 $\pm$ 52.59	265.50	258.85 $\pm$ 35.12	291.92 $\pm$ 36.81	<b>0.015</b> <sup>§</sup>
Median (Min–Max)							
<b>Jitter*</b>	0.80 $\pm$ 0.46	0.79 $\pm$ 1.22	0.56 $\pm$ 0.26	0.74	0.77 $\pm$ 0.31	0.50 $\pm$ 0.15	0.930
Median (Min–Max)							
<b>Shimmer*</b>	6.09 $\pm$ 3.37	4.74 $\pm$ 5.97	5.33 $\pm$ 4.02	5.14	4.28 $\pm$ 1.29	4.24 $\pm$ 2.02	0.898
Median (Min–Max)							
<b>Mean HNR†</b>	16.57 $\pm$ 4.98	18.67 $\pm$ 4.93	18.78 $\pm$ 5.03	16.05	15.07 $\pm$ 5.67	19.11 $\pm$ 3.56	0.368
Mean $\pm$ SD							
<b>Base/Width‡</b>	3.91 $\pm$ 0.63	4.08 $\pm$ 0.90	4.14 $\pm$ 0.76	2.38	3.01 $\pm$ 0.68	3.47 $\pm$ 1.23	<b>0.005</b> <sup>§</sup>
Median (Min–Max)							
<b>Base/Vocal cord*</b>	0.22 $\pm$ 0.05	0.20 $\pm$ 0.05	0.22 $\pm$ 0.05	0.17	0.17 $\pm$ 0.03	0.21 $\pm$ 0.04	0.110
Median (Min–Max)							
<b>Posterior glottic gap ratio*</b>	0.00 $\pm$ 0.00	0.78 $\pm$ 2.52	3.26 $\pm$ 2.60	0.00	1.66 $\pm$ 2.09	4.45 $\pm$ 3.18	<b>0.001</b> <sup>§</sup>
Median (Min–Max)							

Notes: There is no standard deviation in the values because there is only 1 person in group 1d.

\* Kruskal-Wallis *H* test.

† One-way analysis of variance.

‡ 0.05.

§ Difference is statistically significant.

**TABLE 4.**  
**Evaluation of the Correlations of Ratios With Voice Analysis Measurements**

	Mean F0		Jitter		Shimmer		Mean HNR	
	r	P	r	P	r	P	r	P
Base/Width	-0.240	<b>0.042*</b>	-0.305	<b>0.009*</b>	-0.224	0.059	0.324	<b>0.034*</b>
Base/Vocal cord	-0.347	<b>0.023*</b>	0.115	0.334	0.196	0.098	-0.099	0.408
Posterior glottic gap ratio	0.183	0.125	-0.207	0.082	-0.128	0.282	0.162	0.174

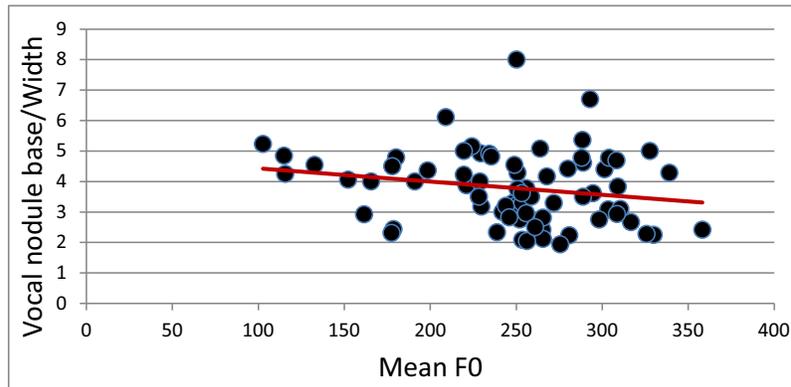
Notes: Spearman correlation test;  $\alpha$ : 0.05.  
 \* Correlation is statistically significant.

and the base/width and posterior glottic gap ratio ( $P = 0.005$ ,  $P = 0.001$ , respectively) (Table 3).

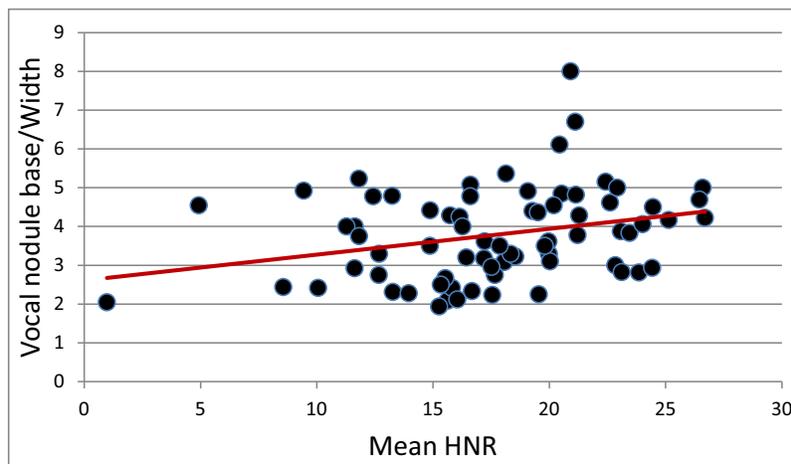
When the effect of these ratios on the voice analysis data was examined, a statistically significant negative correlation was determined between the mean F0 value and the nodule base/width ratio (Table 4). As the vocal nodule base/width ratio increased, the mean F0 value decreased ( $P = 0.042$ ,  $r = -0.240$ ) (Figure 2). A statistically significant negative correlation was determined between the mean F0 value and the base/vocal cord ratio ( $P = 0.023$ ,  $r = -0.347$ ). A statistically significant negative correlation was determined between jitter (%) and vocal nodule base/width ( $P = 0.009$ ,  $r = -0.305$ ). The

effect of the ratios on the shimmer (%) value was not found to be statistically significant. A statistically significant positive correlation was determined between mean HNR and vocal nodule base/width ( $P = 0.034$ ,  $r = 0.324$ ) (Figure 3). According to the discriminant analysis findings, the classifying variables (base/width, base/vocal cord, posterior glottic gap) of the Shah et al scale were found to correctly classify at the rate of 73.6% (Table 5, Figure 4).

The patients were grouped according to gender and age groups of 3–5, 6–9, 10–13, and 14–17 years. A statistically significant difference was observed between the voice analysis measurements and the base/vocal cord ratios ( $P = 0.001$ ) (Tables 6, 7).



**FIGURE 2.** The negative correlation curve between the mean F0 value and the right-side vocal nodule base/width ratio.



**FIGURE 3.** The statistically significant, negative correlation curve between the mean HNR value and the left-side vocal nodule base/width ratio.

**TABLE 5.**  
The Degree of Classification of the Shah et al Scale Classifying Variables According to the Findings of the Discriminant Analysis

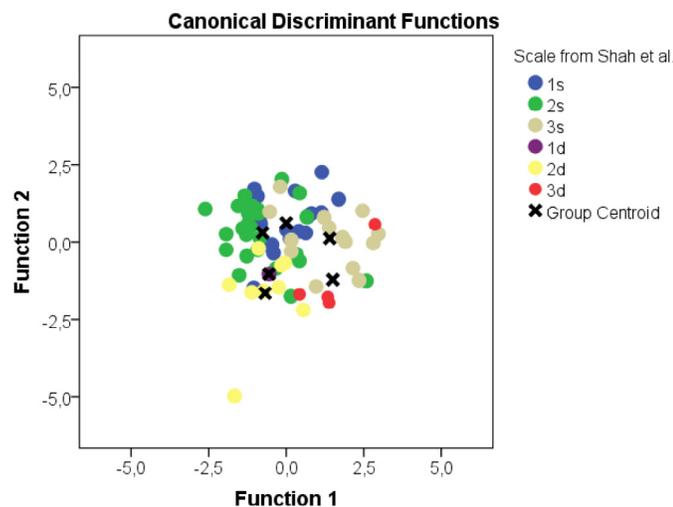
Validated Scale	Predicted Group Membership												Total	
	1s		2s		3s		1d		2d		3d			
	n	%	n	%	n	%	n	%	n	%	n	%	n	%
1s	18	94.7	1	5.3	0	0.0	0	0.0	0	0.0	0	0.0	19	100.0
2s	6	24.0	16	64.0	1	4.0	0	0.0	1	4.0	1	4.0	25	100.0
3s	2	14.3	2	14.3	8	57.1	0	0.0	0	0.0	2	14.3	14	100.0
1d	0	0.0	0	0.0	0	0.0	0	0.0	1	100.0	0	0.0	1	100.0
2d	1	11.1	1	11.1	0	0.0	0	0.0	7	77.8	0	0.0	9	100.0
3d	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	4	100.0	4	100.0

Notes:  $\alpha$ : 0.05; discriminant analysis; correct classification: 73.6%.

Classifier variables: nodule base/width, nodule base/vocal cord, posterior glottic gap ratio, mean F0, jitter, shimmer, mean HNR, age.

## DISCUSSION

It was attempted to develop an objective grading system for use by experienced and inexperienced clinicians on patients with PVNs. When the ratios made of the dimensions of the modules



**FIGURE 4.** Discriminant analysis of the Shah et al scale.

were evaluated with voice analysis, the mean F0 and jitter values were determined to be affected by the width and the mean HNR value by the base of the nodule.

There have been previous studies related to vocal cord pathology, which have used biomedical engineering techniques.<sup>9-12</sup> In those studies, classification has been made according to the blood vessels in the vocal fold surface in pathologies such as vocal cord polyp, nodule, or sulcus.<sup>9,10</sup> In a study by Verikas et al, vocal cords were classified in three groups as nodular, diffuse, or healthy from the color, tissue, and geometric shape properties revealed on vocal cord imaging.<sup>11</sup> In another study, Hilbert analyzed vibration impairments in the vocal cords using transformation.<sup>12,13</sup> Lohscheller et al obtained vocal cord images with a Toy high-speed video-laryngostroboscope device and applied glottis segmentation using the semiautomatic dividing Regional Growth algorithm.<sup>14</sup> In these previous studies, the techniques have been used more for the differentiation of vocal cord pathologies. There have been no studies that have measured ratios and examined the correlation with voice acoustic measurements. Therefore, this study is the first such example.

Vocal fold nodules in childhood are the most common cause of dysphonia. The reason for chronic voice disorders is vocal fold nodules in 38%–78% of cases.<sup>15</sup> The evaluation of funda-

**TABLE 6.**  
Evaluation of the Voice Analysis and Ratios According to Age

	Age (y)				P
	3–5	6–9	10–13	14–17	
	Mean $\pm$ SD	Mean $\pm$ SD	Mean $\pm$ SD	Mean $\pm$ SD	
Mean F0	211.98 $\pm$ 84.59	226.39 $\pm$ 77.60	260.54 $\pm$ 49.46	244.15 $\pm$ 50.72	0.265
Jitter	0.46 $\pm$ 0.36	1.47 $\pm$ 2.14	0.69 $\pm$ 0.26	0.58 $\pm$ 0.31	<b>0.031*</b>
Shimmer	4.00 $\pm$ 3.63	8.84 $\pm$ 10.11	4.78 $\pm$ 2.49	4.59 $\pm$ 3.14	0.080
Mean HNR	18.35 $\pm$ 8.01	16.27 $\pm$ 6.54	17.51 $\pm$ 5.22	18.21 $\pm$ 4.35	0.803
Base width	3.59 $\pm$ 0.35	4.03 $\pm$ 0.87	3.80 $\pm$ 0.91	3.89 $\pm$ 0.88	0.889
Base vocal fold	0.33 $\pm$ 0.06	0.22 $\pm$ 0.04	0.20 $\pm$ 0.05	0.20 $\pm$ 0.04	<b>0.002*</b>
Posterior glottic gap ratio	3.27 $\pm$ 4.63	1.43 $\pm$ 2.44	1.29 $\pm$ 2.31	1.28 $\pm$ 2.67	0.752

Notes: analysis of variance;  $\alpha$ : 0.05.

\* Difference is statistically significant.

**TABLE 7.**  
Correlations of the Ratios According to Age

Age (y)		Base Width		Base Vocal Fold		Posterior Glottic Gap Ratio	
		r	P	r	P	r	P
6–9	Mean F0	–0.141	0.738	–0.348	0.399	0.308	0.458
	Jitter	0.319	0.442	0.865	<b>0.006*</b>	–0.321	0.438
	Shimmer	0.325	0.432	0.906	<b>0.002*</b>	–0.228	0.587
	Mean HNR	–0.171	0.685	–0.725	<b>0.042*</b>	0.279	0.503
10–13	Mean F0	–0.218	0.222	–0.224	0.209	0.273	0.124
	Jitter	–0.323	0.066	–0.089	0.622	–0.079	0.663
	Shimmer	–0.022	0.905	0.254	0.153	–0.141	0.435
	Mean HNR	0.239	0.180	–0.008	0.963	–0.003	0.987
14–17	Mean F0	–0.154	0.424	0.016	0.936	0.096	0.621
	Jitter	–0.102	0.600	0.203	0.290	–0.174	0.367
	Shimmer	–0.221	0.249	0.197	0.306	–0.129	0.506
	Mean HNR	0.235	0.221	–0.181	0.346	0.184	0.340

Notes: Pearson correlation test;  $\alpha$ : 0.05.

\* Correlation is statistically significant.

mental frequencies is important as they are the reason for the sharp drop in values seen in male patients over the age of 12 years. In the current study, the mean age of the patients was 11.3 years. In the literature, vocal fold nodules have been reported to be seen more in young males.<sup>4,16,17</sup> Consistent with those reports in the literature, PVNs were observed more in males in the current study. No statistically significant difference was observed between the male and female children with respect to the mean F0, jitter, and shimmer values. Neither was any statistically significant difference determined between the male and female children with respect to the ratios calculated (base/width, base/vocal cord, posterior glottic gap ratio).

The patients in the current study were grouped according to age and gender. The vast majority of the males were in the 10–13 years age group and the majority of the females were in the 14–17 years age group. Shah et al grouped 600 pediatric patients with vocal nodules in the age ranges of 0–2, 3–5, 6–9, 10–13 and 14–19 years. Large dimension nodules were generally found to be in the 3–10 years age group.<sup>16</sup> In the current study, the base/vocal ratio was evaluated as  $0.33 \pm 0.06$  in the 3–5 years age range and this ratio was observed to be greater in comparison with the other age groups. The reason for this is that with increasing age, there is an increase in the size of the vocal cord, which reduces the ratio. When the age groups were examined, a positive correlation between the jitter and the shimmer base/vocal cord ratio and a negative correlation with mean HNR was seen in the 6–9 years age group. Thus, the increases in the nodule base affected the voice more in this period, causing an increase in jitter and shimmer.

In a study of pediatric patients by Shah et al, a correlation was found between vocal cord nodule size and the severity of hoarseness, breathiness, straining, and aphonia, and nodule size was also found to be correlated with larynx hyperfunction.<sup>16</sup> In another study by Shah et al, no statistically significant correlation was found between nodule size and changes in pitch in the objective and subjective voice measurements of patients with vocal nodules.<sup>18</sup> A total of 145 patients with PVN were evaluated

according to the Shah et al. scale in a study by Nuss et al, and PVN size was determined to increase roughness, strain, pitch, and loudness values of perceptual measurements. However, no relationship was found between the perceptual measurements and the nodule contours (sessile or discrete).<sup>6</sup> According to the mathematical calculation made from the ratios, the mean F0 value in the negative correlation of mean F0 with nodule base/width ratio was affected more by an increase in the width of the nodule. In jitter, again an increase in the width value caused an increase in the jitter value. In contrast, the mean HNR value was determined to be more affected by the base value. In previous studies, there has been no objective comparison between the evaluated nodule size with the base and width measurements of the nodule was made. Therefore, this is the first comparative study made in this area. Nevertheless, longitudinal cross-sectional studies to be applied in the future could provide more useful information about the changes in these ratios with correction by treatment of voice analysis. The current study can be considered to be of guidance in this area.

According to the Shah et al scale, wide-based nodules, in other words, vocal nodules with base length more than twice the width, are classified as sessile nodules and those that are less as discrete nodules.<sup>1,5</sup> In the current study, this ratio was calculated as  $3.82 \pm 0.80$ . The highest mean F0 value in the current study was found in the 3d grade of the Shah et al scale. In the classification of patients according to the Shah et al scale, as the grade of the groups increases, the mean F0 value also increases, and this difference was evaluated as statistically significant. However, the difference in the changes in the jitter, shimmer, and mean HNR values with an increase in the grades of the Shah et al scale classification was not evaluated as statistically significant.

### Study limitations

As the dimensions of the larynx were not equal in the patients and the distance of the vocal fold from the oropharynx was not completely equal in all patients during the videolaryngoscopy, ratios rather than measurements were given in the study. It could

be recommended that in future studies, by equalizing these distances, evaluation of the size of the vocal cord nodules and the area they cover rather than ratios could be made, and the effect of this on the voice could be examined. Furthermore, measurement of these ratios after pediatric nodule treatment and an evaluation of the relationships with voice analysis could extend the scope of future studies.

In the current classification systems, pediatric nodules that affect the voice are classified by the size of nodule, whereas in this study, apart from nodule size, evaluation was made to determine whether the height of the nodule or the dimensions of the base of the nodule affected the voice.

### CONCLUSION

In this study, new ratios were determined in patients with PVN, in collaboration with the biomedical engineering department. The results of the study determined that in the voice acoustic analysis, the mean F0 was more affected by the width of the nodule, mean HNR was affected by the length of the base of the nodule, and jitter (%) was affected by the width of the nodule.

### REFERENCES

1. Shah RK, Feldman HA, Nuss RC. A grading scale for pediatric vocal fold nodules. *Otolaryngol Head Neck Surg.* 2007;136:193–197.
2. Colton RH, Casper JK, Leonard R. Voice problems associated with the pediatric and the geriatric voice. In: *Understanding Voice Problems*. 3rd ed. Philadelphia: Lippincott Williams and Wilkins; 2006:208–209.
3. Nardone HC, Recko T, Huang L, et al. A retrospective review of the progression of pediatric vocal fold nodules. *JAMA Otolaryngol Head Neck Surg.* 2014;140:233–236.
4. Campisi P, Tewfik TL, Manoukian JJ, et al. Computer-assisted voice analysis: establishing a pediatric database. *Arch Otolaryngol Head Neck Surg.* 2002;128:156–160.
5. Nuss RC, Ward J, Recko T, et al. Validation of a pediatric vocal fold nodule rating scale based on digital video images. *Ann Otol Rhinol Laryngol.* 2012;121:1–6.
6. Nuss RC, Ward J, Huang L, et al. Correlation of vocal fold nodule size in children and perceptual assessment of voice quality. *Ann Otol Rhinol Laryngol.* 2010;119:651–655.
7. Boersma P, Weenink D. *Praat: Doing Phonetics by Computer (Version 5.3.53) [Computer Software]*. 2013.
8. Otsu N. A threshold selection method from gray level histogram, IEEE Transaction on system. *Man and Cybernetics.* 1979;9:62–66.
9. Turkmen HI, Karşlıgil ME, Kocak I. Classification of laryngeal disorders based on shape and vascular defects of vocal folds. *Comput Biol Med.* 2015;62:76–85.
10. De Biase NG, Pontes PA. Blood vessels of vocal folds: a videolaryngoscopic study. *Arch Otolaryngol Head Neck Surg.* 2008;134:720–724.
11. Verikas A, Gelzinis A, Bacauskiene M, et al. Towards a computer-aided diagnosis system for vocal cord diseases. *Artif Intell Med.* 2006;36:71–84.
12. Yan Y, Ahmad K, Kunduk M, et al. Analysis of vocal-fold vibrations from high-speed laryngeal images using a Hilbert transform-based methodology. *J Voice.* 2005;19:161–175.
13. Huang NE. *Hilbert-Huang Transform and its applications*. second ed. Danvers, USA: 2014.
14. Lohscheller J, Toy H, Rosanowski F, et al. Clinically evaluated procedure for there construction of vocal fold vibrations from endoscopic digital high-speed videos. *Med Image Anal.* 2007;11:400–413.
15. Von Leden HV. Vocal nodules in children. *Ear Nose Throat J.* 1985;64:473–480.
16. Shah RK, Woodnorth GH, Glynn A, et al. Pediatric vocal nodules: correlation with perceptual voice analysis. *Int J Pediatr Otorhinolaryngol.* 2005;69:903–909.
17. Yano J, Ichimura K, Hoshino T, et al. Personality factors in pathogenesis of polyps and nodules of vocal cords. *Auris Nasus Larynx.* 1982;9:105–110.
18. Shah RK, Engel SH, Choi SS. Relationship between voice quality and vocal nodule size. *Otolaryngol Head Neck Surg.* 2008;139:723–726.