

Quantitative analysis of videokymography in normal and pathological vocal folds: a preliminary study

Cesare Piazza · Stefano Mangili · Francesca Del Bon ·
Francesca Gritti · Claudia Manfredi ·
Piero Nicolai · Giorgio Peretti

Received: 15 April 2011 / Accepted: 14 September 2011 / Published online: 30 September 2011
© Springer-Verlag 2011

Abstract Videokymography (VKG) captures high-speed images of the vocal folds independently of the periodicity of the acoustic signal. The aim of this study was to preliminarily assess a software package that can objectively measure specific parameters of vocal fold vibration. From August 2009 until December 2010, we prospectively evaluated 40 subjects (Group A, 18 normal subjects; Group B, 14 patients with benign lesions of the middle third of the vocal fold, such as polyps and cysts; Group C, 8 patients treated by endoscopic excision of vocal fold benign lesions) by videoendoscopy, videolaryngostroboscopy, and VKG. A VKG camera was coupled to a 70° telescope and video was recorded during phonation. Images were objectively analyzed by a post-processing software tool (VKG-Analyser) with a user-friendly interface developed by our group. Different parameters were considered, including the ratio between the amplitude of the vibration of one vocal fold with respect to the contralateral (R_{amp}), the ratio between the period of one vocal fold vibration and the opposite one (R_{per}), and the ratio between the duration of the open and closed phase within a glottal cycle (R_{oc}). Mean values for R_{amp} , R_{per} , and R_{oc} in Group A were 1.05, 1.04, and 1.35, respectively; in Group B were 1.63, 0.92, and 0.97, respectively; and in Group C were 1.13, 0.91, and 1.85, respectively. Quantitative analysis of videokymograms by the herein presented tool, named

VKG-Analyser, is useful for objective evaluation of the vibratory pattern in normal and pathologic vocal folds. Important future developments of this tool for the study of both physiologic and pathologic patterns of vocal fold vibration can be expected.

Keywords Videokymography · Quantitative analysis · Image processing · Vocal fold vibration · Laryngeal imaging · Laryngoscopy

Introduction

Evaluation of the vibratory pattern of vocal folds is of paramount importance to understand the mechanism of voice production and to obtain accurate differential diagnosis for vocal fold disorders. Moreover, a detailed examination of vocal fold vibration in a quantitative manner could offer more possibilities to individualize treatment strategies, choosing between voice therapy and surgery or, among the latter, between transoral phonosurgical and laryngeal-framework procedures. Furthermore, such objective measures could be of utmost value in order to compare pre- and post-treatment outcomes. A number of different diagnostic tools, such as videolaryngostroboscopy (VLS), high-speed videolaryngoscopy (HSV), and videokymography (VKG), have been introduced in the daily practice with the aim of helping voice specialists in understanding the causes, ideal therapeutic strategies, and functional outcomes of different vocal fold diseases [1, 2].

VLS, even though considered as a first-choice for vocal fold examination, presents some limitations mainly related to the fact that it works properly only in presence of an almost periodic acoustic signal. This hampers the diagnostic possibilities of VLS in evaluating patients with

C. Piazza (✉) · S. Mangili · F. Del Bon · P. Nicolai · G. Peretti
Department of Otorhinolaryngology, Head and Neck Surgery,
University of Brescia, Piazza Spedali Civili 1,
25123 Brescia, Italy
e-mail: ceceplaza@libero.it

F. Gritti · C. Manfredi
Department of Electronics and Telecommunications,
University of Florence, Florence, Italy

extremely hoarse or breathy voices, as well as patients unable to keep a stable pitch during phonation. In light of this, new imaging devices with a capture rate faster than vocal fold vibration have been developed to overcome the problems of irregular voice with fast pitch and register jumps. For example, digital HSV records a full laryngological image at a rate of 2,000 frames/s with a resolution of 512×512 pixels. However, at present, it can store only 4 s of recorded data and generally has poorer image quality as compared with VLS. Moreover, its costs are not negligible [3, 4].

In contrast, VKG offers an excellent image rate (7,812.5 images/s) with a high spatial resolution (768 pixels/line) thanks to the fact that only a single line of the glottic plane is visualized and recorded during examination. VKG has several advantages over full image high-speed systems: lower cost, a smaller amount of data to be stored and processed, and therefore, the possibility to record samples without time constraints. Kymograms can be also extracted from recorded high-speed video files with the advantage to obtain them at different places across the glottic plane (digital kymograms). However, the image quality obtained in this case is worse compared to VKG due to an approximately three-time smaller line resolution and image rate [3, 4]. Recently, a new generation of VKG systems has been introduced that provide simultaneous laryngoscopic and kymographic images for routine clinical vocal fold examination. The scan position of the kymogram is always in the laryngoscopic image, which provides feedback for the operator to easily locate the expected scanning position line [5].

Even if VKG images could be used for evaluation and subsequent analysis of parameters related to the vocal fold vibratory pattern (frequency, amplitude, free margin irregularities, left–right asymmetries, propagation of mucosal waves, duration of the various phases of the glottal cycle, and cycle-to-cycle variability), image processing and computerized automatic quantification of these objective parameters still need further development and refinements.

The aim of this descriptive, preliminary study was to validate a post-processing software tool developed by our group and described in its technical aspects elsewhere [6, 7], provided with a user-friendly interface, capable of measuring and tracking quantitative parameters from VKG examinations. It has been assessed on synthetic images and then clinically applied for the first time here to a group of normal and pathological subjects.

Materials and methods

From August 2009 until December 2010, at the Laryngologic Service of the Department of Otorhinolaryngology,

Head and Neck Surgery, University of Brescia, Italy, we prospectively evaluated 40 subjects (18 males, 22 females; mean age 39.9 years; range 26–56) by VKG as part of a comprehensive laryngeal examination protocol including videoendoscopy and VLS. Eighteen were healthy subjects without any laryngeal complaints or history of previous laryngeal diseases (Group A, control group). Fourteen were affected by benign lesions of the middle third of one vocal fold, which were polyps in seven cases and cysts in seven (Group B). Group C was composed of eight subjects who had been previously treated for benign lesions of the middle third of the vocal fold by phonosurgical procedures at our institution.

The first part of the examination included videoendoscopy and VLS by 70° rigid telescope (Kay Pentax, Lincoln Park, NJ). During VLS, we qualitatively evaluated different parameters such as the presence of mucosal wave, its symmetry, periodicity, and amplitude, glottic competence, and the presence of non-vibrating portions of the vocal folds [8].

All patients were subsequently evaluated by a new generation VKG system (VKG camera 2156, Cymo B.V., Groningen, The Netherlands) that allows simultaneous laryngoscopic and kymographic visualization as described by Qiu and Schutte [5]. Images were obtained with a 70° rigid telescope (Kay Pentax, Lincoln Park, NJ) coupled to a xenon light source (RLS 9100) by one of the two senior authors (G.P. and C.P.). Digital recording was accomplished directly by the VLS station (9200 C). Patients were evaluated in an upright position, without use of local anesthesia. They were asked to utter a sustained vowel /a/ at a modal register, comfortable pitch, and moderate phonatory effort for at least 5 s. We evaluated the videokymograms obtained by VKG-Analyser, a post-processing software tool developed by a member of our group (C.M.) with the aim of measuring and tracking quantitative parameters from VKG examinations [6, 7]. The analysis was focused on videokymograms taken from the middle portion of the glottic plane in Groups A and C, while it was brought through the vocal fold lesion in Group B (Fig. 1).

Specifically, the following parameters were taken into account:

1. R_{amp} , which measures the amplitude of one vocal fold vibration with respect to the opposite one:

$$R_{\text{amp}} = \frac{a_r}{a_l}$$

where a_r and a_l are the amplitudes of right and left vocal folds vibration, respectively (Fig. 2).

2. R_{per} , is defined as the ratio between right (t_r) and left (t_l) vocal fold periods (Fig. 3):

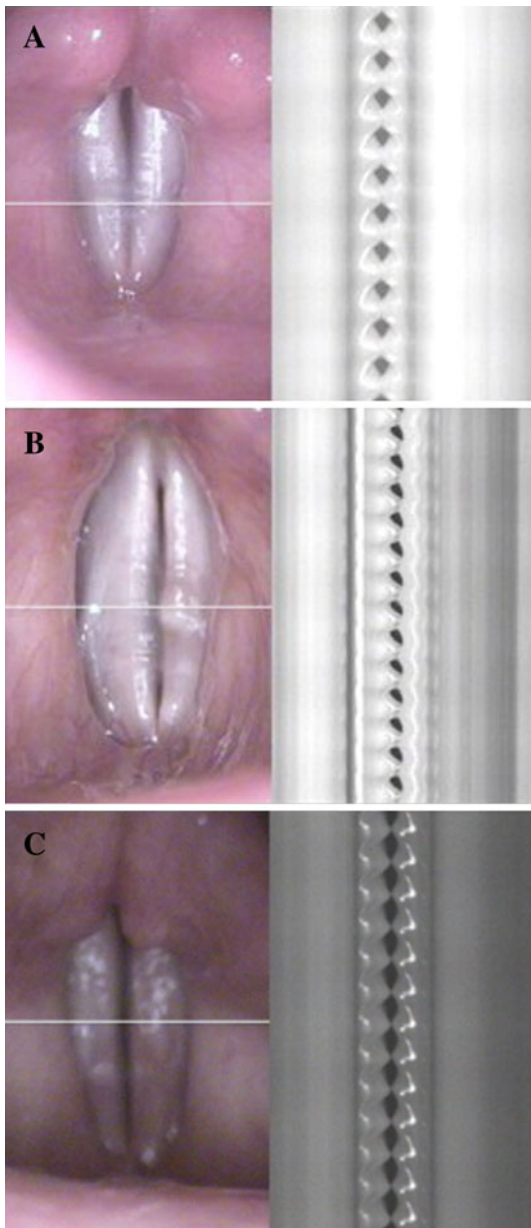
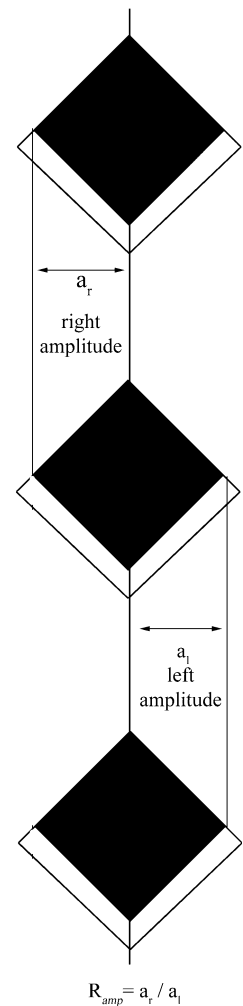


Fig. 1 Videolaryngoscopic and VKG images of three representative patients in the present series: **a** normal subject (Group A); **b** patient with a cyst of the middle third of the left vocal fold (Group B); **c** patient submitted to phonosurgical excision of a cyst of the middle third of the right vocal fold (Group C). Notice the position of the scanning VKG line at the level of the middle third of the vocal fold in each patient

$$R_{per} = \frac{t_r}{t_l}$$

While this is the rule for healthy subjects, in presence of vocal fold lesions, the ratio between healthy and abnormal vocal fold was used for calculation of both R_{amp} and R_{per} . Therefore, in case of disease of the right

Fig. 2 Schematic drawing of a kymogram showing the amplitude of vibration of both vocal folds and related R_{amp} . For explanation see “Materials and methods”



vocal fold, the software was set to calculate the reciprocal of the above ratio.

3. R_{oc} , is the measure of the ratio between the duration of the open (t_o) and closed (t_c) phase within a single glottal cycle:

$$R_{oc} = \frac{t_o}{t_c}$$

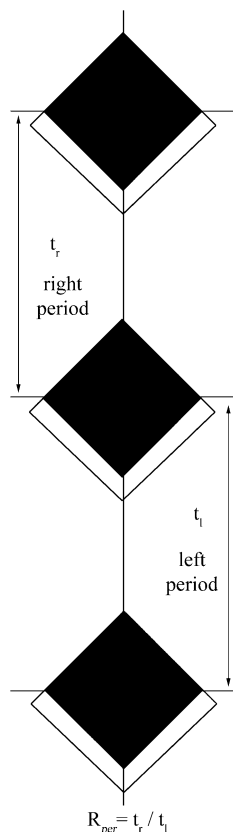
where t_o is the time between the opening point and the following closing point and t_c is the time intervening between the closing point and the following opening point (Fig. 4).

Standard deviation for each one of the objective parameter listed above was also calculated.

Results

None of the normal subjects (Group A) showed any pathological abnormalities in VLS parameters. In Group B, all patients affected by cysts of the middle portion of the vocal

Fig. 3 Schematic drawing of a kymogram showing the period of vibration of both vocal folds and related R_{per} . For explanation see “Materials and methods”

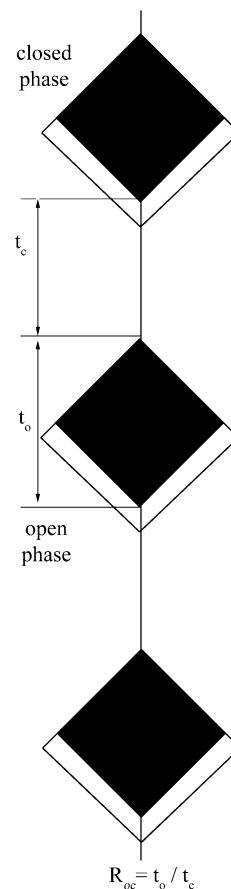


fold showed a reduction in amplitude and symmetry of the mucosal wave, and segmental mucosal wave impairment at the level of the lesion in the affected fold. Incomplete glottic closure was observed in three patients and aperiodicity in only one. In the same group, four of the five subjects affected by polyps presented a reduction of mucosal wave amplitude and symmetry, segmental mucosal wave impairment, aperiodicity of the pathologic vocal fold vibration, and incomplete glottic closure. One patient only showed complete glottic closure. Patients of Group C showed reduction of mucosal wave amplitude and symmetry, and minimal glottic incompetence in all cases. Non-vibrating segment of the vocal fold in the area submitted for endoscopic surgery was noticed in four patients.

The mean values of the quantitative parameters (R_{amp} , R_{per} , and R_{oc}) obtained by VKG-Analyser from VKG examinations of the entire cohort of patients are detailed in Table 1. As expected, in Group A, the values of R_{amp} and R_{per} were both around 1, while R_{oc} was 1.35. This means that, at the level of the middle portion of the healthy vocal folds, the duration of the open phase of the glottal cycle is always longer than that of the closed phase.

In Group B, the value of R_{amp} was found larger than 1 due to the fact that the healthy vocal fold presents wider amplitude of vibration as compared to the affected one. In contrast, the value of R_{per} diminished due to the fact that

Fig. 4 Schematic drawing of a kymogram showing the open and closed phase of a single glottal cycle and related R_{oc} . For explanation see “Materials and methods”



the period of vibration of the affected vocal fold was longer than that of the healthy one. The R_{oc} value was less than the normal value due to the longer duration of the closed phase in the presence of expansile lesions.

In Group C, R_{amp} was larger than in both normal subjects and patients with expansile lesions. This can be explained by the presence of cicatricial stiffness (and ensuing reduction in mucosal wave amplitude) of the operated vocal fold with respect to the healthy one. For the same reason, R_{per} is lower, while R_{oc} is significantly higher than in normal subjects due to minimal iatrogenic glottic incompetence prolonging the open phase of the glottal cycle.

Discussion

Undoubtedly, at present VKG represents one of the most sophisticated methods of endoscopic evaluation of vocal fold vibratory pattern. Its relative low cost, the reduced amount of data to be stored and processed, and the high image rate and spatial resolution certainly represent the most relevant advantages of VKG as compared to full image high-speed systems [1–4]. When comparing VKG to VLS, on the other hand, it clearly represents an adjunctive

Table 1 Mean values and standard deviations (SD) for R_{amp} (ratio of amplitudes), R_{per} (ratio of periods), and R_{oc} (open to closed phase ratio) for Group A (control group), Group B (patients with polyps and cysts), and Group C (patients submitted to phonosurgical excision of expansile lesions)

	R_{amp}	R_{per}	R_{oc}
Group A	1.052 (SD \pm 0.175)	1.04 (SD \pm 0.217)	1.35 (SD \pm 0.447)
Group B	1.639 (SD \pm 0.601)	0.922 (SD \pm 0.246)	0.978 (SD \pm 0.831)
Group C	1.135 (SD \pm 0.568)	0.917 (SD \pm 0.359)	1.851 (SD \pm 1.37)

cost in terms of camera and software to be used, while light source, rigid telescope, and recording system can be shared with those already available in the standard VLS station.

VKG performed by 70° rigid telescope without the use of local anesthesia needs a compliant patient to be properly accomplished. Even though VKG has been experimentally applied to flexible transnasal scopes (unpublished observations), at present its clinical application is somewhat limited by the need of using transoral endoscopes. Moreover, a strict standardization of the endoscopic exam (i.e., position at which the kymogram is recorded and type of phonation produced during recording) is mandatory to minimize subtle alterations of the recorded kymograms, particularly when quantification of these parameters is attempted as in the present study. As a matter of fact, the vibratory pattern is undoubtedly different along the glottis, with maximal amplitudes registered around the middle portion of the vocal folds, and a minimum at both the anterior and posterior ends. For this reason, the VKG scanning line should be ideally placed perpendicular to the glottic axis (with a tolerance of no more than 15°) in order to evaluate the corresponding points of the two vocal folds. The variability of VKG images can also be physiologically influenced by pitch, loudness, and register produced during the examination. Therefore, strict adherence to a protocol, standardizing all the abovementioned variables, is needed to properly interpret VKG findings and to clinically compare different examinations [4].

However, up to now, VKG has been used for objective evaluation of vocal fold vibratory patterns, without quantitative analysis of specific parameters. VKG-Analyser is a post-processing software tool capable of extracting and tracking reliable and reproducible objective parameters that can be easily calculated and used in a clinical setting. It has been provided with a user-friendly interface for real-time analysis of VKG images and also allows storage and retrieval of patient data, even though the present paper is not focused on description of its specific technical features [6, 7].

Even if our results show close correlation between VLS and VKG findings in normal subjects (Group A), kymograms combined with the quantitative analysis offer a complementary tool for investigation of the physiologic

vibratory behavior of vocal folds. This can be certainly influenced by a number of different patient-related physical factors such as vocal fold geometry, tension, pliability, and symmetry, which all need to be taken into account for future VKG quantitative evaluations of “normal” voice production. As a matter of fact, to the best of our knowledge, this is the first time that a precise quantification of the ratio between open and closed phases in the glottal cycle at the level of the middle portion of the vocal fold has been set to 1.35. This, as well as other measurements, will certainly become easily available by a systematic use of VKG-Analyser.

On the other hand, in pathological cases (Group B), where the VLS findings depend on subjective judgement and are thus biased by the examiner, VKG-Analyser can play a significant role, being extremely useful in order to diagnose problems concerning the vibratory pattern. In pathologic subjects, our results show heterogeneous variability of VLS abnormalities, while the mean values of all VKG parameters present a homogeneous trend compared to the control group. The same holds true when dealing with iatrogenic alterations of vocal fold vibration. Whilst VLS can be affected by substantial inter-rater disagreement, VKG-Analyser precisely measures the nature and entity of glottic incompetence and reduced pliability.

The most important flaw of the present study is represented by the relatively small number and heterogeneity of VKG samples analyzed, reflecting the descriptive and preliminary nature of this work, essentially aimed at demonstrating the possibility to apply our software in a clinical setting. This can also explain, at least in part, the large standard deviation encountered, together with the fact that these data were derived from different subjects with different pathologic conditions.

Future investigations should be conducted using case–control studies to minimize the physiological variations of vocal fold vibration between patients. Indeed, one of the most interesting applications of this software could be the possibility to quantitatively analyze pre- and post-treatment changes on a case-by-case basis, thus modulating the different therapeutic options for each individual patient. Up to now, however, we never based our therapeutic choices on the VKG parameters as calculated by the present software, due to the absence of any previous validation. Moreover, another possible application of such a tool could be the storage of pre- and postoperative quantitative VKG data from different patients both for medical-legal and research purposes.

In conclusion, we herein describe VKG-Analyser, a post-processing software tool with a user-friendly interface that is capable of objectively measuring some clinically useful parameters routinely observed in VKG images. Its complementary role in the evaluation of normal and

pathological vocal fold vibratory behavior seems to be extremely promising in the clinical setting.

References

1. Verikas A, Uloza V, Bacauskiene M, Gelzinis A, Kelertas E (2009) Advances in laryngeal imaging. *Eur Arch Otorhinolaryngol* 266:1509–1520
2. Deliyiski DD, Hillman RE (2010) State of the art laryngeal imaging: research and clinical implications. *Curr Opin Otolaryngol Head Neck Surg* 18:147–152
3. Schutte HK, Svec JG, Sram F (1998) First results of clinical application of videokymography. *Laryngoscope* 108:1206–1210
4. Svec JG, Sram F, Schutte HK (2009) Videokymography. In: Fried MP, Ferlito A (eds) *The larynx*. Plural Publishing, San Diego, pp 251–271
5. Qiu Q, Schutte HK (2006) A new generation videokymography for routine clinical vocal fold examination. *Laryngoscope* 116:1824–1828
6. Manfredi C, Peretti G (2006) A new insight into postsurgical objective voice quality evaluation: application to thyroplastic medialization. *IEEE Trans Biomed Eng* 53:442–451
7. Manfredi C, Bocchi L, Bianchi S, Migali L, Cantarella G (2006) Objective vocal fold vibration assessment from videokymographic images. *Biomed Signal Process Control* 1:129–136
8. Hirano M, Bless D (1993) *Videostroboscopic examination of the larynx*. Singular Publishing Group, San Diego