Clinical Practicability of a Newly Developed Real-time Digital Kymographic System

*Jin-Choon Lee, *Soo-Geun Wang, *Eui-Suk Sung, †In-Ho Bae, ‡Seong-Tae Kim, and §Yeon-Woo Lee, *†*Yangsan*, ‡*Naju, and* §*Busan, South Korea*

Summary: Background. A digital kymogram shows real images of vocal fold vibration. However, DKG is difficult to use in clinical practice because the recorded image cannot be seen instantaneously after examination, as considerable encoding time is required to visualize a digital kymogram. In addition, frame-by frame analysis should be implemented to evaluate high-speed videoendoscopy data, but is time- and labor-intensive.

Purpose. The purpose of the study was to validate the clinical practicability of a real-time multislice digital kymographic system developed by the authors. We analyzed the promptness and accuracy of the examination before and after intracordal injections in patients with unilateral vocal fold paralysis.

Methods. To assess the clinical applicability of this system, six patients with unilateral vocal fold paralysis were selected. Real-time DKG was performed before and immediately after intracordal injection. We observed changes in the digital kymogram after the intracordal injection.

Results. Using this system, 10 scanning lines and up to five vertical pixel row could be obtained in real time, and the maximum acquisition time for the DKG image was 10 seconds. A digital kymogram of the patients could be instantaneously acquired, and whether the intracordal injection was appropriate or not.

Conclusion. This article is the first validation study after the development of the real-time multislice digital kymographic system. Our system may be a promising tool in clinical practice for immediate assessment of the vibratory patterns of the vocal cords. More research is necessary for further clinical validation.

Key Words: Real-time–Digital kymography–Intracordal injection–Vocal cord–Paralysis.

INTRODUCTION

Observation of vocal fold vibration in patients with dysphonia is essential for accurate diagnosis and evaluation before and after treatment. Laryngeal videostroboscopy has been used primarily in clinical practice. However, the images from videostroboscopy are illusory, as they are collected from different cycles. In addition, it is impossible to examine vocal cord vibration in cases of very irregular vocal fold vibration or short sustained vowel phonation. Videokymography[1] addresses some of the main shortcomings of videostroboscopy, and is displayed in real time, but an important limitation remains that only a single line of the laryngeal image is scanned and displayed.

Laryngeal high-speed videoendoscopy (HSV)[2] can be supplanted to solve the limitations of videostroboscopy. HSV provides more accurate information on vocal fold vibration. However, HSV has not been widely used in clinical practice because of practical limitations, such as the considerable time necessary for encoding and recording, the extremely large storage necessary for archiving, and the intense concentration required to analyze HSV images.

Recently, many researchers introduced postprocessing analytic methods from HSV data such as digital kymography (DKG),[3] glottal area waveform,[4] glottal width waveform,[5] and kymographic edge detection.[6] These methods could provide information on the characteristics of vocal fold vibrations. However, these methods, with the exception of digital kymography, might need complicated software applications, and the results are sometimes difficult to interpret. The DKG images are extracted from the HSV images and show the real vibratory image of the vocal folds. However, there are some disadvantages, such as a considerable waiting time before kymographic visualization and a recording duration limited to a few seconds.

With advanced computer performance and software technology, we developed a real-time multislice DKG system to overcome the limitations of the conventional DKG system. The purpose of this study was to demonstrate the feasibility of the real-time, multislice DKG system in clinical practice. We determined the promptness and accuracy of the analysis of unilateral vocal fold paralysis before and after intracordal injection.

MATERIALS AND METHODS

Subjects

One normophonic male (36 years old) participated in this study for the verification of the capacity of our system. Six patients with unilateral vocal fold paralysis (mean age, 57.83 years; range, 43~76 years; female:male = 1:1) were selected.

Instruments

To assess the vibratory pattern of the vocal cords, we used a black and white complementary metal-oxide-semiconductor camera with a global shutter (USC-700MF, U-medical, Busan,

Accepted for publication October 31, 2017.

From the *Department of Otorhinolaryngology-Head and Neck Surgery, Pusan National University School of Medicine, Yangsan, Gyeongsangnam-do, South Korea; †Department of Otorhinolaryngology-Head and Neck Surgery, Pusan National University Yangsan Hospital, Yangsan, Gyeongsangnam-do, South Korea; ‡Department of Speech-Language Pathology, Dongshin University, Naju, Jeollanam-do, South Korea; and the §Department of Otorhinolaryngology-Head and Neck Surgery, Pusan National University Hospital, Busan, South Korea.

Address correspondence and reprint requests to Soo-Geun Wang, Department of Otorhinolaryngology-Head and Neck Surgery, Pusan National University School of Medicine, 20, Geumo-ro, Beomeo-ri, Mulgeum-eup, Yangsan-si, Gyeongsangnam-do, South Korea. E-mail: entwangsg@gmail.com

Journal of Voice, Vol. 33, No. 3, pp. 346–351
0892-1997

https://doi.org/10.1016/j.jvoice.2017.10.024

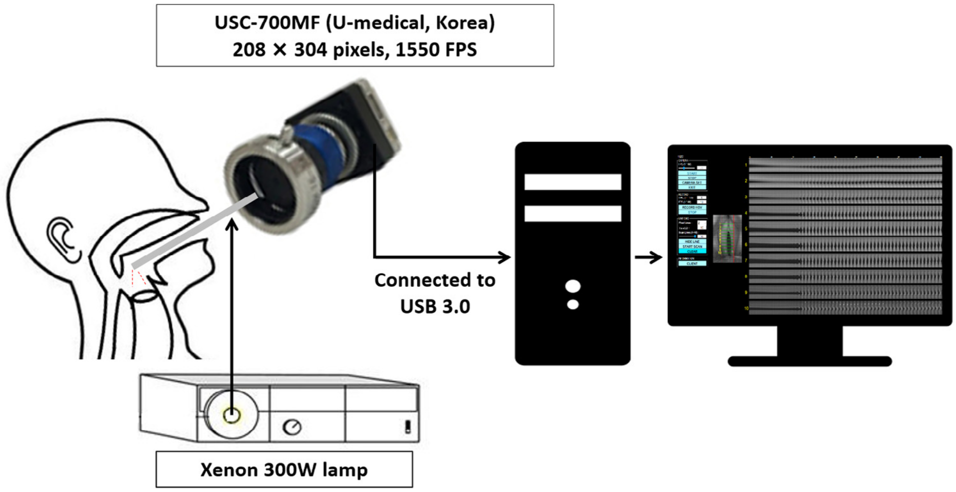

FIGURE 1. Schematic illustration of a real-time multislice DKG system.

Korea). The spatial resolution was 208 × 304 pixels and the frame rate per second was about 1500. However, the frame varies in real-time measurement depending on the computer performance. We used a 4-mm diameter, 70-degree, rigid laryngoscope (8700 CKA, Storz, Tuttlingen, Germany) connected to a zoom coupler (f = 16–34 mm, MGB, Eschbach, Germany). A 300-W xenon light source (NOVA 300, Storz, Tuttlingen, Germany) was used (Figure 1).

To create the DKG images, we developed a software program using Visual Studio Integrated Development Environment. We used an algorithm that concatenated a single vertical pixel row from each frame of the high-speed images to generate a digital kymogram the same as Tigges et al.[3] The scanning line means the perpendicular line to glottal axis. Our newly developed multiline, real-time digital kymography can be selected up to 10 scanning lines. Vertical pixel row means the minimum unit of monitor's height, and we controlled arbitrarily from one to five pixel rows.

The minimum requirements for the computer system are window 10 (64-bit), Intel Core i7, Ram 8 gigabyte, Full-HD (1920 × 1080) resolution, and USB 3.0 port. The image information of the camera is transferred to the program through USB 3.0. The camera used in this study acquires an image of about 1500 frames/s at a resolution of 208 × 304. Because the video data of about 90.45 megabytes/s (208 × 304 × 1500 × 8 bit = 758,784,000 bit/s) are transferred from the camera to the computer program, for this purpose, USB 3.0 which can transmit up to 5 gigabytes/s (625 megabytes/s) should be used.

The configuration of the user interface is shown in Figure 2 (see video clip 1). The function of the multislice DKG system could be controlled by the buttons displayed in Figure 2A. Laryngeal images for navigation are displayed in Figure 2B. The multislice DKG image is shown in Figure 2C. DKG imaging was achieved as still image files (JPEG format), and time information of each frame is saved separately as text file (txt format).

Procedures

The HSV camera connected with a laryngoscope is inserted through the oral cavity to examine the vocal cords. Ten scanning lines, a recording time of 10 seconds, and two vertical pixel rows were set, and the examination was performed after the scanning line position was adjusted in the navigation laryngeal image.

All participants were instructed to phonate a sustained vowel /e/ or /i/ with comfortable pitch and intensity during DKG examination. When recording starts using the foot switch or start button, digital kymogram images can be obtained in quasi-real time (see video clip 2). The navigation laryngoscope video image continues to be displayed after 10 seconds.

All patients with vocal cord paralysis were checked by real-time DKG before and immediately after intracordal injection. To evaluate whether the amount and site of the injection were appropriate, we classified the state of the intracordal injection into three groups: under-, ideal, and overcorrection of the paralyzed cord. We also identified superficial injections, indicated by the absence of vocal cord vibration because the filler injected superficially.

Adding to the functionality of our system, we designed HSV that can be recorded separately from DKG. Using Kang's DKG conversion program from HSV playbacks, temporal change of DKG was made by superimposing one to five vertical pixel rows (Figure 3).[7]

RESULTS

Validation of real-time DKG in a vocally healthy participant

We tested the performance of real-time DKG in a normophonic participant. The results are summarized in Table 1. Up to 10 scanning lines can be selected, and DKG can be represented differently according to vertical pixel rows from one to five. The vertical pixel rows mean the number of vertical pixels to form one scanning line.

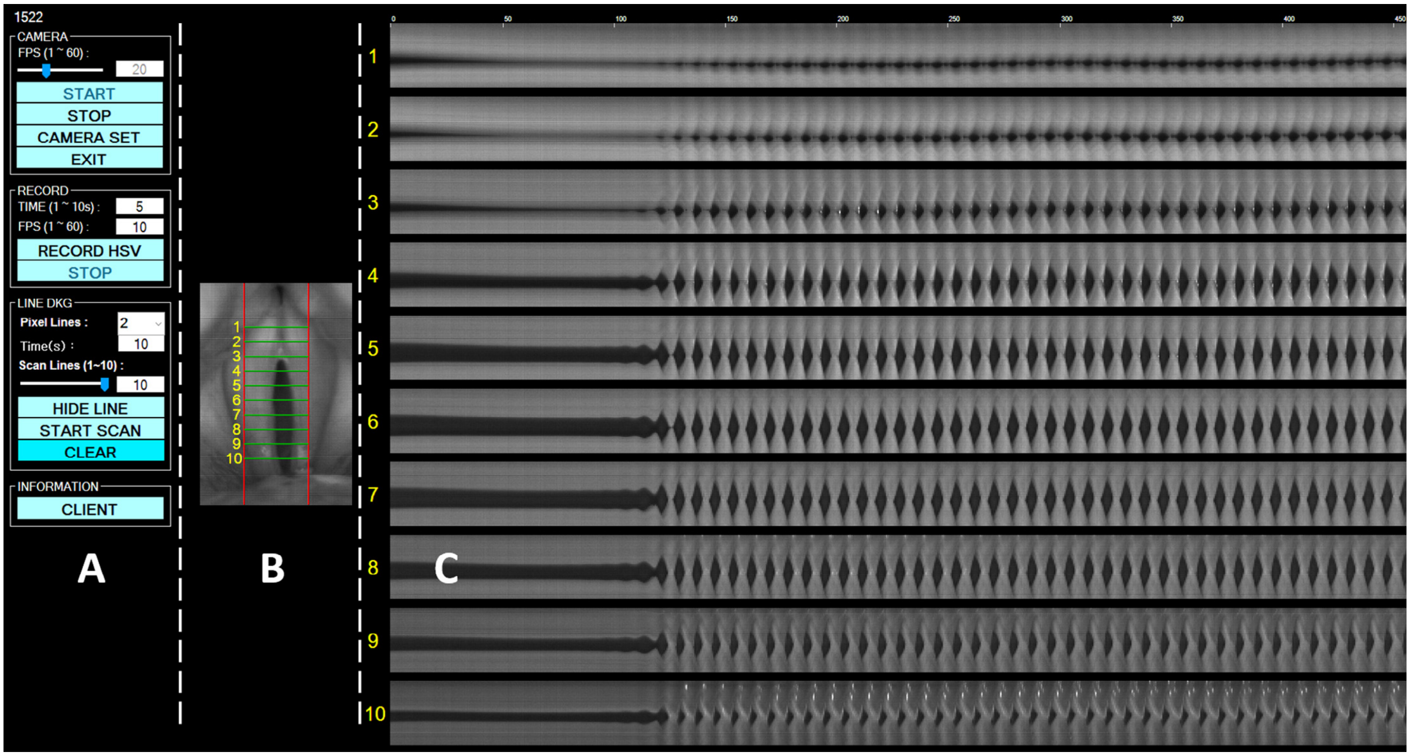

FIGURE 2. Configuration of the user interface of a real-time multislice DKG system. **A.** Option button for camera setting and recording. **B.** Navigation of the laryngeal image. **C.** Output of the multislice DKG. The timeline shows above the DKG image.

According to the change in the number of vertical pixel rows in the same phonation, different time resolution of digital kymograms could be obtained (Figure 3). Five kymograpic images were obtained separately from the same phonation. As the number of vertical pixel rows increases, only the time resolution increases at the same frame rate (about 1500 frames/s). The number of pixel row applied equally to all scanning lines (see video clip 1).

Figure 4 represents the output of our real-time DKG in a normophonic male. Using a laryngeal navigation image, DKG images were displayed in the horizontal direction for a maximum of 10 seconds. The time domain at intervals of 50 ms was displayed above the DKG images.

Feasibility of clinical use for the assessment of unilateral vocal cord paralysis

To determine the feasibility of clinical use of our real-time DKG, we evaluated the vibratory pattern of the vocal cords before and immediately after the intracordal injection of the filler. Parts A and B of Figure 5 represent undercorrection of the paralyzed vocal cords. Parts C, D, and E of Figure 5 represent ideal correction of the paralyzed cord. Part F reveals overcorrection due to a superficial injection, and no vibratory pattern was noted in the left cord. The real-time DKG system was useful for rapid recognition of the postinjection state.

DISCUSSION

Laryngeal videostroboscopy is being replaced by HSV because of the several aforementioned drawbacks. HSV can capture at least 10–20 frames per vibratory cycle and present the true intracycle vibratory characteristics of the vocal folds independent of cycle-to-cycle periodicity.

Deliyski[8] reported that a high-speed camera with a speed of at least 4000 frames/s can track the details of mucosal wave propagation. In this situation, the encoding time increased because of the large file size and the recording duration was limited to a few seconds. Considerable time and concentration are required for frame-by-frame analysis of the HSV data.[9] For these reasons, the commercial HSV system has not been frequently used in clinical practice.

Recently, many researchers confirmed that postprocessing methods such as glottal area waveform,[4] laryngotopography,[10] and DKG from HSV playbacks can provide useful clinical information for assessment of the characteristics of vocal fold vibrations.[11,12] The evaluation of vocal fold vibrations was much more effective when using the functional postprocessing methods with HSV than on analysis of the HSV alone. Glottal area waveform and laryngotopography are considered to be useful approaches for the assessment of spatial characteristics. However, DKG is used widely to analyze vibratory characteristics such as phase symmetry between the left and the right and anterior and posterior, amplitude asymmetry, mucosal wave of the vocal folds, and voice onset and offset.[3]

If DKG is intended to be clinically practicable, it must be displayed quickly after HSV examination. Our system can obtain multiple DKG images and display them in a static form directly during videolaryngoscopy examination. This is possible through

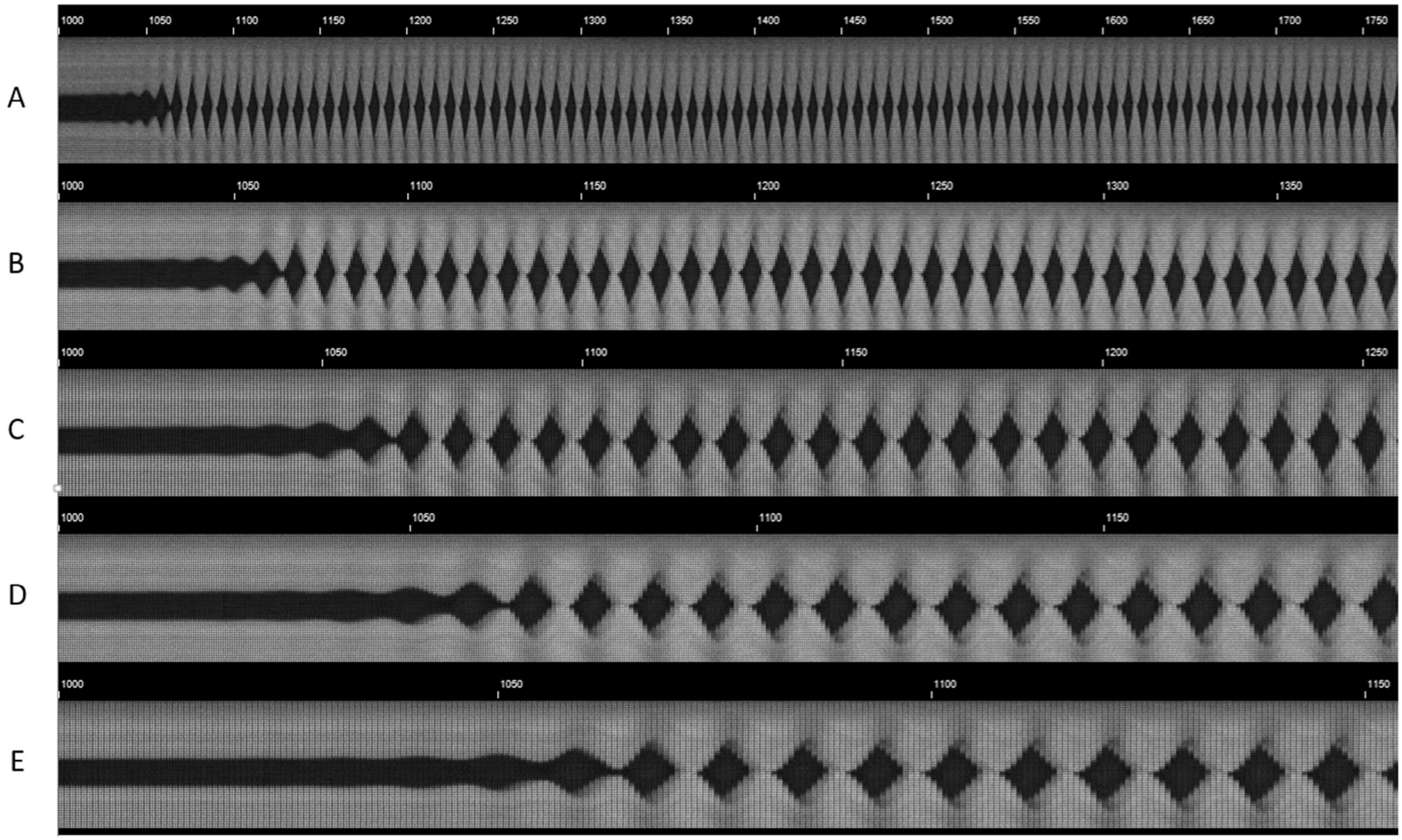

FIGURE 3. Digital kymogram according to the accumulation of vertical pixel rows. **A.** One pixel row, **B.** Two pixel rows, **C.** Three pixel rows, **D.** Four pixel rows, and **E**. Five pixel rows. As the number of vertical pixel row was changed, different time resolution in the digital kymogram could be observed. The fundamental frequency (F_0, 120 kHz) could be calculated in the DKG of every pixel row.

the improvement of the software algorithm and computer performance.

We can reduce the encoding and recording times significantly. Our system, and videokymography developed by Švec and Schutte[1] are methods by which real-time kymography can be conducted. The latter is displayed on a TV monitor in AVI format and contains 40 ms of information and can connect kymogram of unlimited time if necessary.[13,14]

The characteristics of our system are as follows. The scanning line can include a maximum of 10 lines, because the examiner cannot change the position of the scanning lines during the examination. The evaluation of the vocal cord vibrations is sufficient even if one or two scanning lines are outside of the region of vibrations. Because the number of pixel row can be adjusted from one to five, vocal cord vibration can be represented with various time resolutions. For men with a low fundamental frequency, the number of pixel row was increased slightly from one to two or three, and for women with a higher fundamental frequency, it was decreased slightly from three to two or one for a more accurate evaluation. As shown in Figure 3B, we can easily evaluate the vocal fold mucosa through an image reconstructed with two vertical pixel rows in general.

To evaluate the clinical practicability of our real-time DKG system, we compared the pre- and postinjection DKG findings of patients with unilateral vocal cord paralysis. If undercorrection because of insufficient intracordal injection is suspected, additional reinjection is necessary. For this reason, rapid identification

TABLE 1.
Technical Specifications of the Quasi-Real-time Multislice Digital Kymography System

Time to display images	Quasi-Real-time
Total acquisition time	Maximum of 10 s
Number of scanning lines used to create image*	Maximum of 10 lines
Number of vertical pixel row†	1~5 pixel rows
File format of archived digital kymography images	Stored automatically in the designated folder in JPEG and txt format

* Scanning line refers to perpendicular line to glottal axis, which can be selected by examiner's preference.
† Vertical pixel row means the minimum unit of monitor's height.

FIGURE 4. Sequential images of a multislice digital kymogram (36-year-old normophonic man). **A.** Navigation laryngeal image. **B.** A digital kymogram with 10 scanning lines. The vertical pixel row was applied to two pixel rows.

of vocal fold vibration after intracordal injection is helpful. Among several examination modalities, it is difficult to accurately observe vocal cord vibrations with videolaryngoscopy, laryngeal videostroboscopy requires attachment of EGG or laryngeal microphone, and conventional DKG, including HSV, is time-consuming to store and convert data, making it difficult to immediately assess the appropriateness of intracordal injection. To overcome these limitations, quasi-real-time DKG could be helpful. The advantage of DKG is that it can analyze the vibration characteristics of each vocal cord quantitatively compared with laryngoscopy or laryngeal videostroboscopy. With DKG, it is possible even for beginners to easily analyze and accurately examine vocal cord vibrations of subjects with severe dysphonia, which is not possible with a stroboscopic examination.

In our system, mucosal wave could not be traced clearly because of low frame rate, and smooth edge of kymogram cannot be expressed when the number of vertical pixel row increases. Thus, aliasing phenomenon will appear in the kymogram. These problems can be solved when the performance of computers and frame rates of cameras are improved.

CONCLUSIONS

This is the first report of a real-time multislice DKG system being used to analyze normophonic and pathological voice. Using our system, clinicians can quickly observe and evaluate the vibratory pattern of the vocal folds. Our system may be a promising tool in clinical practice, but more research is necessary for further clinical validation. In addition, we expect our system to be included in hoarseness guidelines[15] and to be a routine examination modality equivalent to videolaryngoscopy.

Acknowledgment

This work was supported by a Clinical Research grant in 2016 from Pusan National University Hospital.

SUPPLEMENTARY DATA

Supplementary data to this article can be found online at doi:10.1016/j.jvoice.2017.10.024.

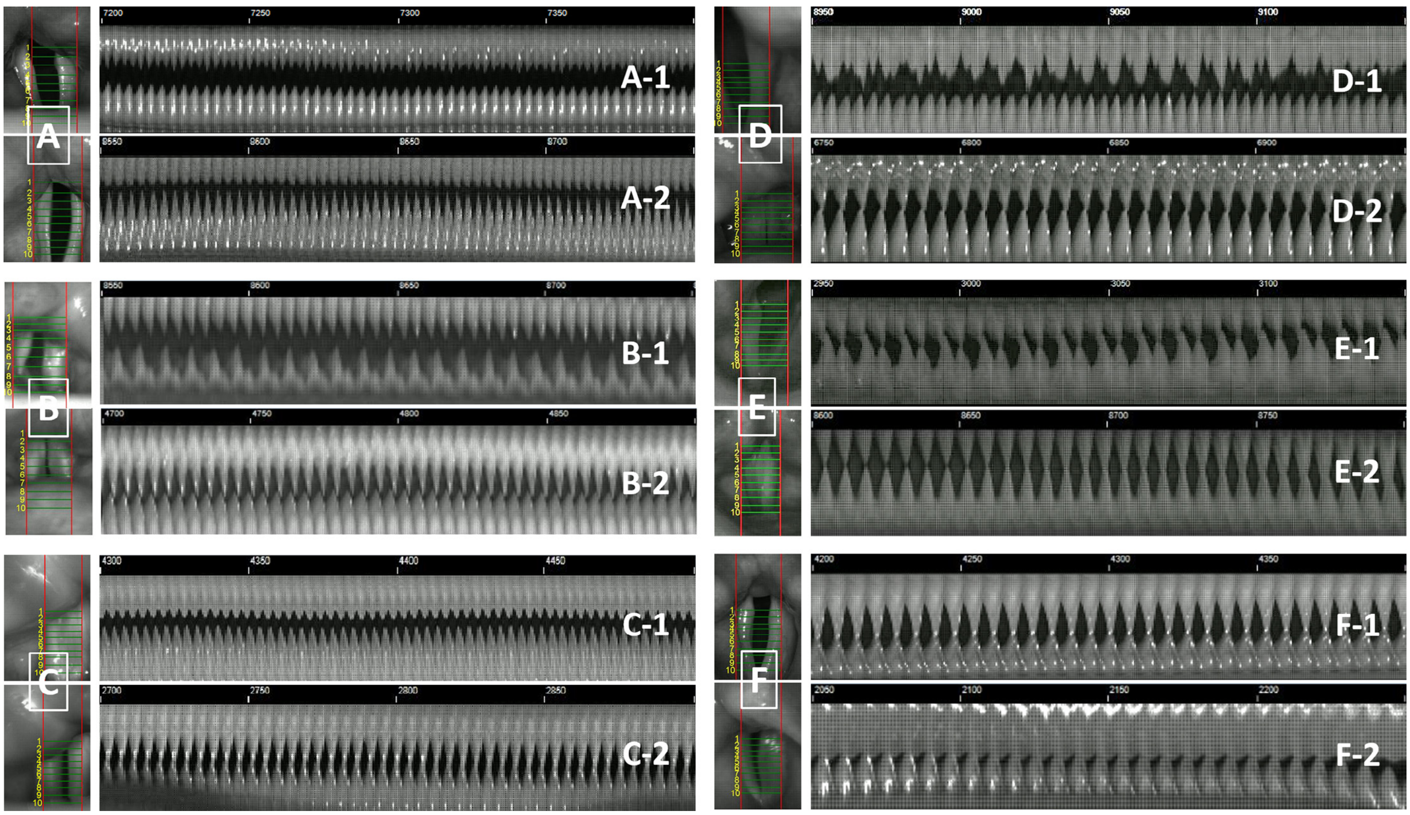

FIGURE 5. Vibratory pattern of vocal cords before and immediately after intracordal injection in vocal fold paralysis cases. **A.** Left side palsy (F/65), preinjection (A-1), postinjection (A-2). **B.** Right side palsy (M/52), preinjection (B-1), postinjection (B-2). **C.** Left side palsy (F/43), preinjection (C-1), postinjection (C-2). **D.** Left side palsy (F/61), preinjection (D-1), postinjection (D-2). **E.** Right side palsy (M/76), preinjection (E-1), postinjection (E-2). **F.** Left side palsy (M/50), preinjection (F-1), postinjection (F-2).

REFERENCES

1. Švec JG, Schutte HK. Videokymography: high-speed line scanning of vocal fold vibration. *J Voice* 1996;10:201–205.
2. Hirose H. High-speed digital imaging of vocal fold vibration. *Acta Otolaryngol* 1988;105(S458):151–153.
3. Tigges M, Wittenberg T Mergell P, Eysholdt U. Imaging of vocal fold vibration by digital multi-plane kymography. *Comput Med Imaging Graph* 1999;23:323–330.
4. Noordzij JP, Woo P. Glottal area waveform analysis of benign vocal fold lesions before and after surgery. *Ann Otol Rhinol Laryngol* 2000;109:441–446.
5. Lohscheller J, Švec JG Döllinger M. Vocal fold vibration amplitude, open quotient, speed quotient and their variability along glottal length: kymographic data from normal subjects. *Logoped Phoniatr Vocol* 2013;38:182–192.
6. Chen W, Woo P, Murry T. Spectral analysis of digital kymography in normal adult vocal fold vibration. *J Voice* 2014;28:356–361.
7. Kang DH, Wang SG, Park HJ, et al. Real-time simultaneous DKG and 2D DKG using high-speed digital camera. *J Voice* 2017;31:247–e1.
8. Deliyski D. Laryngeal high-speed videoendoscopy. In: Kendall KA, Leonard RJ, eds. *Laryngeal evaluation: indirect laryngoscopy to high-speed digital imaging*. New York, NY: Thieme Medical Publishers; 2010:243–270.
9. Yamauchi A, Imagawa H, Yokonishi H, et al. Evaluation of vocal fold vibration with an assessment form for high-speed digital imaging: comparative study between healthy young and elderly subjects. *J Voice* 2012;26:742–750.
10. Yamauchi A, Imagawa H, Sakakibara KI, et al. Phase difference of vocally healthy subjects in high-speed digital imaging analyzed with laryngotopography. *J Voice* 2013;27:39–45.
11. Yamauchi A, Yokonishi H, Imagawa H, et al. Yamasoba T. Quantitative analysis of digital videokymography: a preliminary study on age-and gender-related difference of vocal fold vibration in normal speakers. *J Voice* 2015;29:109–119.
12. Zhang Y, Huang N, Calawerts W, et al. Quantifying the subharmonic mucosal wave in excised larynges via digital kymography. *J Voice* 2017;31:123–e7.
13. Qiu Q, Schutte HK. A new generation videokymography for routine clinical vocal-fold examination. *Laryngoscope* 2006;116:1824–1828.
14. Qiu Q, Schutte HK. Real-time kymographic imaging for visualizing human vocal-fold vibratory function. *Rev Sci Instrum* 2007;78:024302.
15. Schwartz SR, Cohen SM, Dailey SH, et al. Clinical practice guideline: hoarseness (dysphonia). *Otolaryngol Head Neck Surg* 2009;141(suppl 2):S1–S31.