

Short Commentary

Open Access, Volume 5

Computerized video image applications in combination with voice-related biomarkers in medical diagnostics of artistic voices

Mette Pedersen, MD, PhD

Ear-Nose-Throat Specialist, Head and Neck Surgeon Consultant, Denmark.

***Corresponding Author: Mette Pedersen, MD, PhD**

Ear-Nose-Throat specialist, Head and Neck Surgeon
Consultant, The Medical Centre Østergade 18 1100
Copenhagen, Denmark.
Email: m.f.pedersen@dadlnet.dk

Received: Jun 21, 2025

Accepted: Jul 24, 2025

Published: Jul 31, 2025

Archived: www.jclinmedimages.org

Copyright: © Pedersen M (2025).

Abstract

Medical voice diagnostics in artistic medicine has been based on tradition. In the 21st century, with musicals and a half-classical rock and pop tradition there are demands on new kinds of diagnostics. A computerized setup of 16 possible video images (four in use) at a time with a switcher for several cameras, and a Sony screen combined with acoustical analyses, is presented. The switcher Atem Mini Pro 2018 and the sound mixer with 18 channels, CQ-18T Allen and Health hardware, together with a MacBook for computerized measures, were in use at a rock concert. Voice-related acoustical parameters were used. A combination of all voice-related biomarkers was suggested. The setup makes it easier to get a good validation of artistic presenters. Focus was on the synergy of mimics and expression, movements, rhythm, and singing technique - combined with voice online. The combination of video images with voice analysis gives a better background for medical diagnostics than voice-related biomarkers alone and makes comparison between clinical centres easier. In the future AI can be of great help.

Introduction

Most medical diagnostics of voices in artistic medicine are based on tradition. The history hereof is great when referring to singers' satisfaction, but mostly described sporadically in scientific literature, as is the case for general aspects of singing. It was shown that in a simple search with the keyword voice pedagogy on Google Scholar, more than 15.000 hits were found. The question discussed by the authors was how to evaluate the quality of the papers [1]. The development of singing tradition from ancient verbally delivered methods till the paradigm shift of the invention of the larynx mirror by Garcia [2] to the physiological understanding of voice [3] has changed the diagnostics of the artistic voice. Endoscopy with AI of the vocal folds is a big research area [4]. This is also the case for the developing voice [5]. But voice diagnostics up to the 21st century with musical and half-classical rock and pop tradition once more demands new medical diagnostics [6,7].

Due to the historical development of voice traditions, demands have changed for the medical diagnostics of voice [8]. There have been changes in the tradition of voice diagnostics and documentation. It has been shown that the definition of hoarseness is weak, and no randomized controlled trials (RCTs) thereof were found [9]. There was no unanimity on measuring methods till recently, in an updated Delphi questionnaire where a consensus of voice measures was made [10].

The consensus between the European Laryngological Society and the Union of European Phoniaticians was a great step forward, generating suggestions for voice-related biomarkers in a recent book [11]. In the book voice-related biomarkers of Voice Handicap Index (VHI), the GRBAS test of voice perception, Maximum Phonation Time (MPT) of air flow, and basic acoustic parameters (Fundamental frequency (F0)), Jitter, Shimmer, and Harmonics to Noise Ratio (HNR)) are discussed. The measures

can be used routinely in medical voice diagnostics – as suggested in combination with updated computed video imaging.

The article aimed to show the technical combination of computer video imaging and the voices of singers, and how the combination with voice-related biomarkers can function in medical practice.

Methods and materials

A Switcher was used for 16 computed images on a Sony Screen (Atem Mini Pro 2018). The switcher is a hardware piece used to switch between and transition among multiple video sources, commonly used for live productions. In our case, the evaluation of the setup was carried out at a rock concert where the singers were fulfilling a 3-year singing education. The setup was handled by one of the singing teachers (CW) and discussed with a medical doctor (MP).

The MacBook was used to run the live-production software, managing streaming, recording, graphics overlays, timecode logging, and to archive synchronized audio–video files, sent directly to YouTube, but also usable for post-performance analysis.

Two cameras were connected, giving 4 Camera video images on the screen. A sound mixer was attached with 18 channels out of which four were in use for the acoustic measures referred to as part of voice-related biomarkers (CQ-18, T Allen & Heath). Supplementary voice-related biomarkers can be attached to the measures.

The handling of the setup included securing video images that presented the synergy of mimic (especially smiling) and all-around verbal expression, visually shown movements, and rhythm. Based on the changes in the presentation hereof, relevant camera shifts were handled online, also taking the singing technique into account. The acoustic measures were followed and handled online at the same time, with some chosen standard results given on the screen. Space was possible for the VHI, GRBAS test, and MPT.

Results

Computer Imaging with four cameras makes it possible to switch to the best presentation part, but also to the worst.

The 4 cameras make it easier to make a clear validation of the presentation, also from a medical point of view.

The singer's presentation can be filmed in detail, for among others, posture, jaw position, and the other focuses mentioned above:

Synergy of mimics and expression, visualisation of movements in detail, rhythm, universally and in detail, personal expression and evaluation of the singing technique.

The computerized technical setup can be routinely combined with the voice-related parameters, some online and some stored [11]. Artistic evaluation is much more difficult, but some standards for the future can be considered based on the concert of 36 singers with eventual scores (1-5?) like synergy of mimic and expression, rhythm, personal expression – of song texts, and musical technique.

The results are of a computerized standard that enables inter-collegial discussion and AI analysis with deep learning [8,11].

Discussion

The discussion by Herbst and Meyer [1] of the findings of the Google Scholar search in the library system, must give rise to consideration of new aspects of the medical forms of voice diagnostics. RCT evaluations are very difficult to make [9], but the approach must be considered. Without (meta-)analyses of RCTs, no good and valid conclusions about medical voice diagnostics can be made, as is the case in other medical fields. The use of Delphi questionnaires is better than nothing [4,10] and should be made, eventually related to the methods mentioned above.

The voice-related biomarkers are based on a consensus after 14 webinars published recently [11]. It does not include images, nor of singers either. Separate standards for the medical evaluation of artistic performances that include video images, and voice-related biomarkers can add helpful perspectives to the treatment of singers.

An aspect is the use of AI. There was agreement based on a Delphi questionnaire that AI is useful in larynx endoscopy [4]. Probably, this will also be the case for video images of performers. A perspective is the combination of all the parameters.

It was shown that the acoustical measures in AI validations of Parkinson patients for 10 years, were insufficiently carried out for use in clinical diagnostics [11]. But over time, the development of AI with updated AI models will probably be useful. For now, visual scores combined with (parts of) voice-related biomarkers might be considered a step forward, also enough for comparison between clinical centres and for research, at best with RCTs.

Conclusion

The possible combination of high-quality computerized video images, and voice-related biomarkers for clinical voice diagnostics was presented. Better clinical voice diagnostics is hereby a perspective, eventually using scores for synergy of e.g., mimic (smile) and expression, movements, rhythm, and singing technique. Future AI models may make combined results quicker and even more exact.

Acknowledgment: Claes Wegener recorded the images/videos and is thanked for discussing his ideas at a concert of the Complete Vocal Institute, Copenhagen, Denmark.

References

1. Herbst CT, Meyer D. Critical appraisal of Science-Informed Voice Pedagogy Publications. *J Singing*. 2024; 80: 563–573.
2. Garcia M. A complete treatise on the art of singing: complete and unabridged. Internet Archive. 1805–1906.
3. Titze IR. Principles of Voice Production. USA: Prentice Hall. 1994.
4. Kim YE, et al. AAO-HNSF 2024 Annual Meeting & OTO EXPO. 2024.
5. Pedersen M. Normal development of voice in childhood. Berlin, Heidelberg, New York: Springer Publishers. 2024.

-
6. Sadolin C. Complete vocal technique. Copenhagen: Complete Vocal Institute. 2021.
 7. Estill J, et al. The Estill Voice Model: Theory & Translation. 2017.
 8. Pedersen M. Voice-related Biomarkers in Singing: Suggested Clinical Measurements Standards of Voice Complaints, Listeners' Evaluation, Acoustical Measurement and Airflow. *J Clin Med Res.* 2025; 6: 1–6.
 9. Pedersen M, McGlashan J. Surgical versus non-surgical interventions for vocal cord nodules. *Cochrane Database Syst Rev.* 2012: CD001934.
 10. Lechien JR, et al. Consensus for voice quality assessment in clinical practice: guidelines of the European Laryngological Society and Union of the European Phoniaticians. *Eur Arch Otorhinolaryngol.* 2023; 280: 5459–5473.
 11. Pedersen M, Camesasca V, Nashaat NH, Hernández-Villoria R, Das S. *Voice Related Biomarkers.* Berlin, Heidelberg, New York: Springer International Publishing AG. 2025.