Videomics may reduce intra-and inter-rater variability in the assessment of upper aerodigestive tract endoscopy

Paderno A MD PhD
Ear-nose-throat specialist, head and neck surgeon
Humanitas University, Via Rita Levi Montalcini, 4, 20072 Pieve Emanuele MI, Italy

Pedersen M MD PhD
Ear-nose-throat specialist, head and neck surgeon
The Medical Center, Østergade 18, 1100 Copenhagen, Denmark

Conventional endoscopic evaluation remains a highly subjective process, often dependent on the specialist's experience, intuition, and individual interpretation of visual cues. This subjectivity leads to substantial variability between different observers and even within the same observer over time, as highlighted in multiple studies[1,2]. The application of videomics, through advanced artificial intelligence (AI) models, holds the potential to significantly reduce intra- and inter-rater variability in the assessment of upper aerodigestive tract (UAT) endoscopy.

There are several reasons for the variability. One of them is that the UAT includes many aspects of e.g. anatomy and reduced vision of details of the vocal folds. In a material of 15.732 highspeed films, half of the films were insufficient highspeed films for presenting the whole area of the vocal folds, the material was based on an analysis with a stiff scope through the mouth [3]. The results of the application of AI are discussed in several reviews [4,5].

Videomics may offer a solution to the inter- and intra-variability of endoscopy. As presented in many studies the specificity and sensitivity are high in AI-related endoscopies of the upper digestive tract [6]. Computer vision models, such as convolutional neural networks (CNNs), have demonstrated the ability to provide more consistent and objective assessments. However, the performance of these models is inherently linked to the quality and diversity of the training data, pointing towards the importance of selecting representative frames for diagnostic purposes. [7,8] By standardizing the interpretation of endoscopic images, AI-driven assistance systems can guide physicians toward more uniform decisions, thereby minimizing intra-rater variability. This consistency is achieved by delivering constant, unbiased indications that are unaffected by the external factors that typically influence human judgment, such as fatigue or cognitive biases.

Moreover, AI models have the potential to bridge the gap between specialists with varying levels of training and experience. By providing an objective, probabilistic classification of endoscopic findings, these systems can reduce interpersonal differences, leading to a more standardized assessment across different practitioners.

A prospective randomized control trial (RCT) comparing videomics with traditional endoscopies might document better results than one that only includes traditional RCTs. It is needed to compare RCTs that include videomics before the clinical use of videomics. In other fields e.g. dermatology there is a consent that sensitivity and specificity must be >90 before AI can be used in the clinical diagnostics of malignant melanomas. This approach also has to be discussed in the videomics of the upper aerodigestive tract endoscopy. Till now specificity and sensitivity in image-based ear-nose-throat disorders is mostly <90%. A review of the application of AI in office laryngoscopy as late as 2022 does not discuss the aspect of sensitivity and specificity [4].

In the future supplementary software can help secure better and stable videomics of the upper aerodigestive tract. A possibility is to specifically add measurements of the vocal folds to guarantee a correct analysis [9]. Many other aspects can be discussed.

Disclosures

Competing interests: None.

ORCID Id

Alberto Paderno https://orcid.org/0000-0002-1621-2142
Mette Pedersen https://orcid.org/0000-0003-3566-8513

formaterede: Italiensk (Italien)

Kommenterede [AP1]: I would add:

Ferrari M, Mularoni F, Taboni S, Crosetti E, Pessina C, Carobbio ALC, Montalto N, Marchi F, Vural A, Paderno A, Caprioli S, Gaudioso P, Fermi M, Rigoni F, Saccardo T, Contro G, Ruaro A, Lo Manto A, Varago C, Baldovin M, Bandolin L, Filauro M, Sampieri C, Missale F, Ioppi A, Carta F, Ramanzin M, Ravanelli M, Maiolo V, Bertotto I, Del Bon F, Lancini D, Mariani C, Marrosu V, Tatti M, Cağlı S, Yüce I, Gündoğ M, Dogan S, Anile G, Gottardi C, Busato F, Vallin A, Gennarini F, Bossi P, Ghi MG, Lionello M, Zanoletti E, Marioni G, Maroldi R, Mattioli F, Puxeddu R, Bertolin A, Presutti L, Piazza C, Succo G, Peretti G, Nicolai P. How reliable is assessment of true vocal cord-arytenoid unit mobility in patients affected by laryngeal cancer? a multi-institutional study on 366 patients from the ARYFIX collaborative group. Oral Oncol. 2024 May;152:106744. doi: 10.1016/j.oraloncology.2024.106744. Epub 2024 Mar 22. Erratum in: Oral Oncol. 2024 Aug;155:106887. doi: 10.1016/j.oraloncology.2024.106887. PMID: 38520756.

formaterede: Italiensk (Italien)

References

1. Conde B, Martins N, Brandão M, Pimenta AC, Winck JC. Upper Airway Video Endoscopy: Assessment of the response to positive pressure ventilation and mechanical in-exsufflation. *Pulmonology*, 2019;25(5), 299–304. https://doi.org/10.1016/j.pulmoe.2019.02.008
2. Tuijn S, Janssens F, Robben P, van den Bergh H. Reducing interrater variability and improving health care: a meta-analytical review. *Journal of evaluation in clinical practice*, 2012;18(4), 887–895. https://doi.org/10.1111/j.1365-2753.2011.01705.x
3. Pedersen M, Larsen CF. Accuracy of Laryngoscopy for Quantitative Vocal Fold Analysis in Combination with AI, A Cohort Study of Manual Artefacts. Sch J Oto 6(3)-2021. SJO. MS.ID.000237. DOI: 10.32474/SJO.2021.06.000237
4. Yao P, Usman M, Chen YH, German A, Andreadis K, Mages K, Rameau A. Applications of Artificial Intelligence to Office Laryngoscopy: A Scoping Review. *The Laryngoscope*, 2022;132(10), 1993–2016. https://doi.org/10.1002/lary.29886
5. Wu Q, Wang X, Liang G, Luo X, Zhou M, Deng H, Zhang Y, Huang X, Yang Q. Advances in Image-Based Artificial Intelligence in Otorhinolaryngology-Head and Neck Surgery: A Systematic Review. *Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery*, 2023;169(5), 1132–1142. https://doi.org/10.1002/ohn.391
6. Cho WK, Lee YJ, Joo HA, Jeong IS, Choi Y, Nam SY, Kim SY, Choi SH. Diagnostic Accuracies of Laryngeal Diseases Using a Convolutional Neural Network-Based Image Classification System. *The Laryngoscope*, 2021;131(11), 2558–2566. https://doi.org/10.1002/lary.29595
7. Yao P, Witte D, Gimonet H, German A, Andreadis K, Cheng M, Sulica L, Elemento O, Barnes J, Rameau A. Automatic classification of informative laryngoscopic images using deep learning. *Laryngoscope investigative otolaryngology*, 2022;7(2), 460–466. https://doi.org/10.1002/lio2.754
8. Li C, Jing B, Ke L, Li B, Xia W, He C, Qian C, Zhao C, Mai H, Chen M, Cao K, Mo H, Guo L, Chen Q, Tang L, Qiu W, Yu Y, Liang H, Huang X, Liu G, et al. Development and validation of an endoscopic images-based deep learning model for detection with nasopharyngeal malignancies. *Cancer communications (London, England)*, 2018;38(1), 59. https://doi.org/10.1186/s40880-018-0325-9
9. Pedersen M, Larsen CF, Madsen B, Eeg M. Localization and quantification of glottal gaps on deep learning segmentation of vocal folds. Scientific Reports, 2013;13(1). https://doi.org/10.1038/s41598-023-27980-y