



Validity of tissue homogeneity in confocal laser endomicroscopy on the diagnosis of laryngeal and hypopharyngeal squamous cell carcinoma

Matti Sievert¹ · Marc Aubreville² · Antoniu-Oreste Gostian¹ · Konstantinos Mantsopoulos¹ · Michael Koch¹ · Sarina Katrin Mueller¹ · Markus Eckstein³ · Robin Rupp¹ · Florian Stelzle⁴ · Nicolai Oetter⁴ · Andreas Maier⁵ · Heinrich Iro¹ · Miguel Goncalves⁶

Received: 24 October 2021 / Accepted: 7 February 2022 / Published online: 28 February 2022
© The Author(s) 2022

Abstract

Purpose Confocal laser endomicroscopy (CLE) allows imaging of the laryngeal mucosa in a thousand-fold magnification. This study analyzes differences in tissue homogeneity between healthy mucosa and squamous cell carcinoma (SCC) via CLE.

Materials and methods We included five SCC patients with planned total laryngectomy in this study between October 2020 and February 2021. We captured CLE scans of the tumor and healthy mucosa. Analysis of image homogeneity to diagnose SCC was performed by measuring the signal intensity in four regions of interest (ROI) in each frame in a total of 60 sequences. Each sequence was assigned to the corresponding histological pattern, derived from hematoxylin and eosin staining. In addition, we recorded the subjective evaluation of seven investigators regarding tissue homogeneity.

Results Out of 3600 images, 1620 (45%) correlated with benign mucosa and 1980 (55%) with SCC. ROIs of benign mucosa and SCC had a mean and standard deviation (SD) of signal intensity of, respectively, 232.1 ± 3.34 and 467.3 ± 9.72 ($P < 0.001$). The mean SD between the four different ROIs was 39.1 ± 1.03 for benign and 101.5 ± 2.6 for SCC frames ($P < 0.001$). In addition, homogeneity yielded a sensitivity and specificity of 81.8% and 86.2%, respectively, regarding the investigator-dependent analysis.

Conclusions SCC shows a significant tissue inhomogeneity in comparison to the healthy epithelium. The results support this feature's importance in identifying malignant mucosa areas during CLE examination. However, the examiner-dependent evaluation emphasizes that homogeneity is a sub-criterion that must be considered in a broad context.

Keywords Confocal laser endomicroscopy · Head and neck cancer · Classification system · Non-invasive histological imaging · Larynx · Pharynx

✉ Matti Sievert
matti.sievert@uk-erlangen.de

¹ Department of Otorhinolaryngology, Head and Neck Surgery, Friedrich-Alexander-Universität Erlangen-Nürnberg, University Hospital, Waldstrasse 1, 91054 Erlangen, Germany

² Technische Hochschule Ingolstadt, Ingolstadt, Germany

³ Institute of Pathology, Friedrich-Alexander-Universität Erlangen-Nürnberg, University Hospital, Erlangen, Germany

⁴ Department of Maxillofacial Surgery, Friedrich-Alexander-Universität Erlangen-Nürnberg, University Hospital, Erlangen, Germany

⁵ Pattern Recognition Laboratory, Computer Science, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

⁶ Department of Otorhinolaryngology, Plastic Head and Neck Surgery, Rheinische Westfälische Technische Hochschule Aachen, University Hospital, Aachen, Germany

Introduction

Squamous cell carcinoma (SCC) is considered responsible for over 90% of all pharyngeal and nearly 100% of the larynx malignancies [1]. In the diagnostic process, microscopical, histopathological assessment of tissue biopsy is currently considered the gold standard. Various non-invasive optical imaging methods have emerged as potential alternatives to invasive tissue biopsy, albeit mostly in experimental settings. The most promising are confocal laser endomicroscopy (CLE) and optical coherence tomography [2–5]. Probe-based CLE enables the visualization of live mucosal tissue at a cellular level and has been particularly established in mucosal diagnostics of the gastrointestinal tract in the last decade [6]. CLE has been broadly studied in identification of dysplasia in Barrett's esophagus, gastric and colorectal polyps [7, 8]. Furthermore, it has been shown to be beneficial due to its dynamic in-vivo examination in the assessment of disease activity and development of dysplasia in inflammatory bowel disease and celiac disease [9, 10]. After initially being intensively studied in gastroenterology, research has expanded to other areas, such as pneumology [11], urology [12], neurosurgery [13], and the head and neck region. This method utilizes a small laser scanning probe (1–3 mm) gently applied to the area of interest. A field of view of up to 600 μm enables a magnification up to 1000 times and the real-time visualization of the superficial tissue architecture, with a defined focus depth of usually 60–350 μm [14]. Using fluorescein to outline the intercellular spaces and visualization of blood vessels [6], CLE provides “real-time” optical biopsies [15]. The images acquired by probe-based CLE reassemble histopathological tangential sections of the mucosa at the defined depth given by the probe. The reliable interpretation and classification in malignant and normal findings are not trivial, require training, and are usually aided by classification systems [16–19]. In these classification systems, capillary aberrations and tissue homogeneity play a significant role. Even though an 80–90% accuracy is reported in earlier publications, examiner dependency is still a major problem to broad implementation of this technique [20–22].

Tissue inhomogeneity in SCC is a component of several classification systems [16–19]. In most cases, this feature appears to be reliable in identifying malignancy; however, to the best of our knowledge, an objective analysis of this parameter did not take place yet. We hypothesize that malignant tissue shows, even in the small area investigated at any given point by probe-based CLE of 240 μm , a quantifiable regional variation that correlates to histology and is significantly different from the healthy epithelium. This study aimed to objectively—examiner

independent—assess the differences in tissue homogeneity between healthy epithelium and SCC in CLE compared to examiner dependent analyses related to histopathological findings.

Materials and methods

Study design

This prospective pilot study was conducted at a tertiary hospital and academic cancer center (Department of otorhinolaryngology, head and neck surgery, Friedrich-Alexander-University of Erlangen-Nuremberg, Erlangen, Germany). The study was approved by the local institutional ethics committee (approval number 60_14 B) and carried out following the declaration of Helsinki. In addition, we obtained written informed consent from all study participants.

Eligibility criteria

A total of five consecutive patients with confirmed SCC of the pharynx or larynx were included in this study. Exclusion criteria were a prior treatment of any head and neck cancer, distant metastasis, radiotherapy in the head and neck area, pregnancy, thyroid dysfunction, minority, severe kidney failure, and allergy to fluorescein.

Confocal laser endomicroscopy system and data acquisition

We performed intraoperative imaging using a GastroFlex probe and a 488 nm Cellvizio™ laser scanning system (Mauna Technologies, Paris, France). The probe has a diameter of 2.6 mm, a penetration depth of 55–65 μm , a field of view of 240 μm , and a resolution of 1 μm . We used 5 ml fluorescein alcon, 10% (Alcon PHARMA, Freiburg, Germany) as an optical imaging agent. Surgery began with elevation of the apron flap, following mobilization of the larynx. The second step was to perform a pharyngotomy with cold instruments to avoid thermocoagulation damage to the mucosa. The installation of the CLE probe followed the exposure of the tumor. Subsequently, 2.5 ml fluorescein alcon 10% was injected intravenously. After around 8–10 min of examination, we applied additional 2.5 ml to increase imaging quality. We collected images of the marginal tumor region and the incision margin in the hypopharyngeal mucosa with the CLE probe. The recorded areas were marked with a suture, or a separate biopsy was performed at the precise image acquisition location. In this way, we could correlate CLE imaging with the gold standard of histopathology. The histopathological assessment followed a standard protocol with

hematoxylin and eosin (H&E) staining. After completing the CLE examination, we performed the tumor resection with a macroscopic safety margin of 1 cm. Thus, our and international treatment standards were not altered or influenced in any way by the application of CLE.

Evaluation of tissue homogeneity

For data processing, we analyzed the sequences using Cellvizio Viewer software 1.6.2. by defining four ROIs with rectangular shapes of $45\ \mu\text{m} \times 45\ \mu\text{m}$ (Figs. 1 and 2) in all four image edges. After fluorescein administration, we visually assessed tissue flooding in the region of interest (ROI) and calculated a kinetic graph of the mean histogram value (Fig. 1). After selecting the sequence with an adequate fluorescein signal, we computed a kinetic graph for each ROI. Finally, we compared four different image sectors in a sequence of 60 images over 5 s. For this, we computed the standard deviation (SD) of all four histogram values in the four ROIs to correlate the variability of signal intensity or gray-values within a sequence (Fig. 2). Furthermore, all sequences were categorized into malignant or benign by the reference standard of histopathological analysis.

In addition, we demonstrated the 60 sequences to seven independent raters in an observer-dependent assessment of tissue homogeneity, blinded to the histology, for dichotomous classification into “homogeneous” and “inhomogeneous” tissue architecture. Four raters (rater 1–4) had prior CLE experience, and three were inexperienced (rater 5–7). We defined surgeons with expertise in the technique as having performed at least 20 CLE cases.

Statistical analysis

The two-tailed *t* test for independent samples was applied to compare each histogram's mean value and SD. For the categorical variables, we used the chi-squared test. A *P* value of less than $P < 0.05$ was considered statistically significant. The inter-rater reliability/agreement was tested using the Fleiss kappa coefficient. We interpreted κ -values according to Landis and Koch [23]. Values of κ between 0.0 and 0.20 are defined as low, between 0.21 and 0.40 as fair, between 0.41 and 0.60 as moderate, between 0.61 and 0.80 as substantial, and between 0.81 and 1.0 almost perfect. We performed statistical analysis using SPSS version 26.0 (IBM SPSS Statistics for Windows, Version 26.0. Armonk, NY, USA).

Results

Patient cohort

Between October 2020 and February 2021, we enrolled five patients (all male; mean age 65.4 years (SD=11.9) to undergo in vivo CLE during planned transcervical tumor resection, concerned the hypopharynx and larynx. In one patient (20%), the tumor mass was located in the larynx. Four patients (80%) additionally presented with involvement of the hypopharyngeal mucosa. The tumor resection was performed via an open approach in each case. Microvascular defect reconstruction was necessary in three cases (60%). Regarding the tumor grading, one patient (20%) was confirmed as having an intermediate grade (G2) and four patients (80%) poor grade differentiation (G3). Patient characteristics, including stage, are presented in Table 1. In all

Table 1 Characteristics of patient cohort

Case No.	Age (years)	Tumor stage	Location	Grade	Surgery	CLE frames (<i>n</i>)	Recording time (seconds)	Selected sequences (<i>n</i>)
1	71	T4a	Larynx Hypopharynx	G3	Total LE, partial pharyngectomy	1468	183	14
2	56	T4a	Larynx Hypopharynx	G3	Total LE, partial pharyngectomy	2204	275	12
3	86	T4a	Larynx	G2	Total LE	2191	273	10
4	61	T2	Larynx Hypopharynx	G3	Total LE, partial pharyngectomy	3311	413	14
5	53	T3	Larynx Hypopharynx	G3	Total LE, partial pharyngectomy	2891	361	10
Total						12,065	1505	60

LE laryngectomy

Table 2 Data analysis by the various raters

Inhomogeneity	Rater 1	Rater 2	Rater 3	Rater 4	Rater 5	Rater 6	Rater 7	All
Healthy (<i>n</i> = 27)	1/27 (3.7%)	13/27 (48.1%)	6/27 (22.2%)	1/27 (3.7%)	4/27 (14.8%)	1/27 (3.7%)	0/27 (0%)	26/189 (14.8%)
SCC (<i>n</i> = 33)	15/33 (45.4%)	32/33 (96.9%)	26/33 (78.7%)	30/33 (90.9%)	28/33 (84.8%)	28/33 (84.8%)	30/33 (90.9%)	189/231 (81.8%)
Sensitivity (95%CI)	59.1% (44.4–72.3)	96.9% (84.6–99.4)	78.7% (62.2–89.3)	90.9% (76.4–96.8)	84.8% (69.0–93.3)	84.8% (69.0–93.3)	90.0% (74.3–96.5)	81.8% (76.3–86.2)
Specificity (95%CI)	93.7% (71.6–98.8)	51.8% (33.9–69.2)	77.7% (59.2–89.3)	96.3% (81.7–99.3)	85.1% (67.5–94.0)	96.3% (81.7–99.3)	100% (87.1–100)	86.2% (80.6–90.4)

All comparisons were significant in the chi-square test ($P < 0.001$)

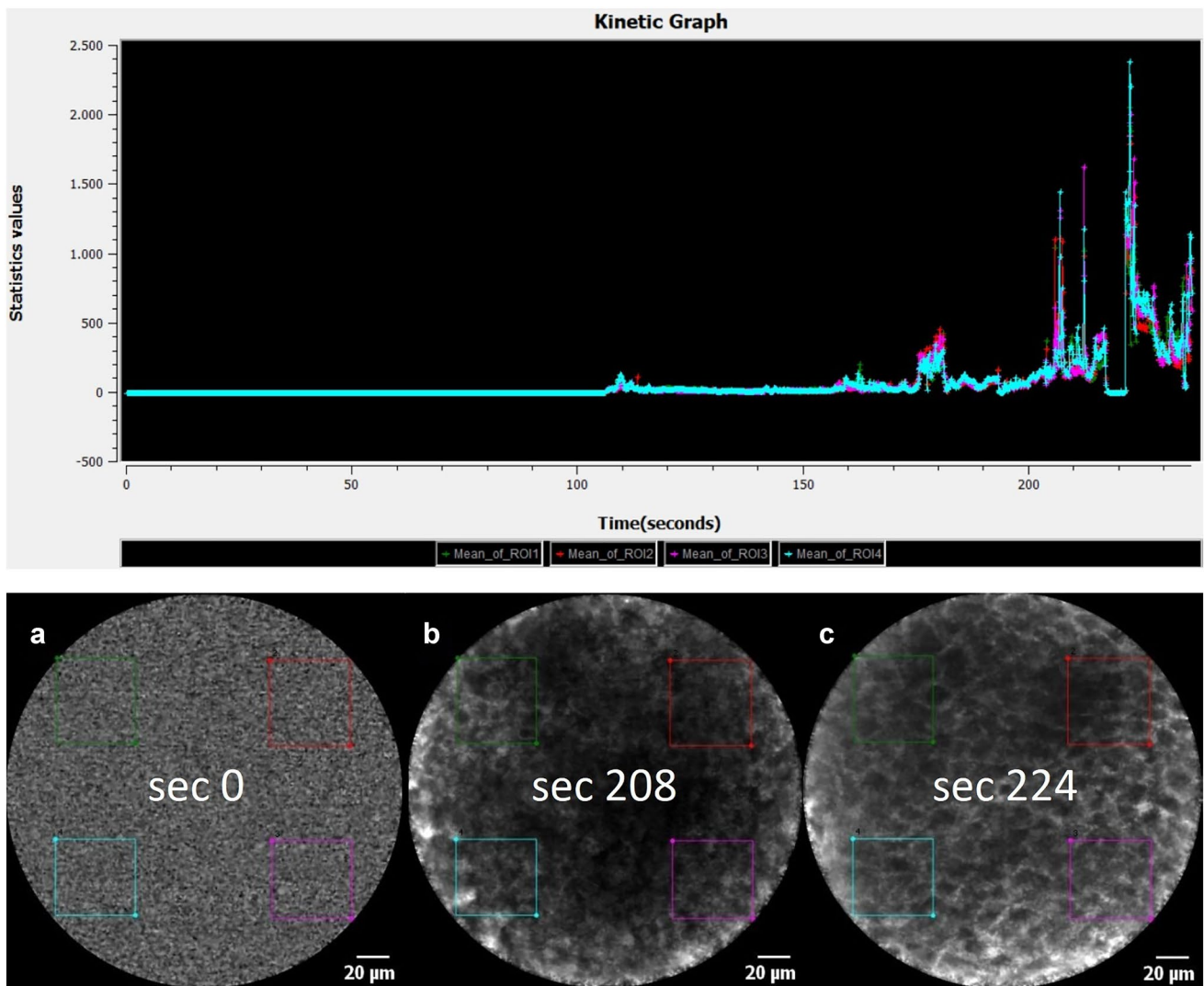


Fig. 1 Kinetic graph shows the flooding of fluorescein after 0 (a), 208 (b), and 224 s (c) after i.v. application of 5 ml, fluorescein alcon, 10%. Acceptable quality is seen after 224 s after initial flooding

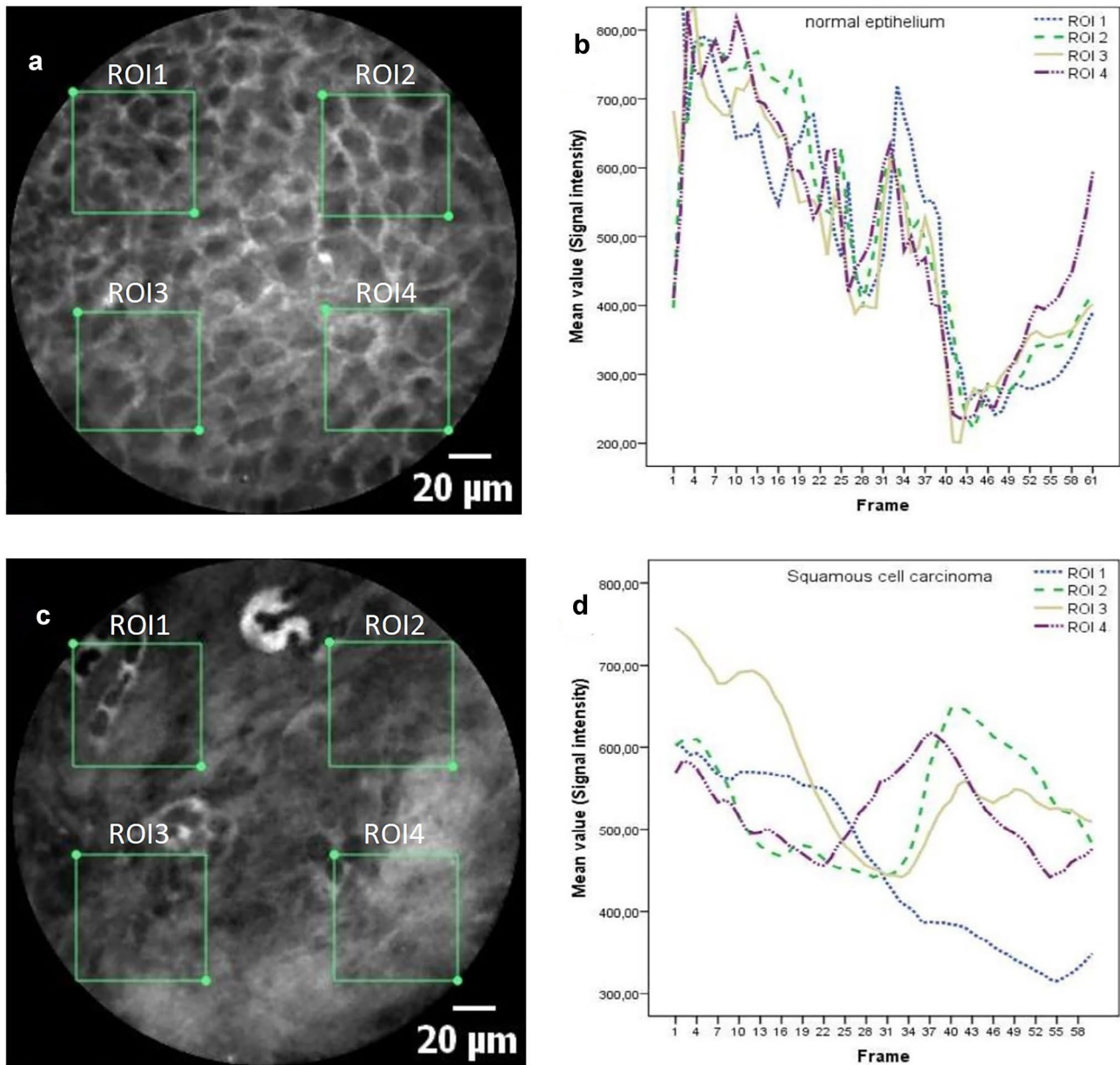


Fig. 2 Healthy epithelium (**a**) and squamous cell carcinoma; SCC (**b**): tissue homogeneity, clear cell borders with regular, similarly sized cells with a small nucleus/cytoplasm ratio, ideally form a honeycomb pattern, are typical features of benign mucosal tissue (**a**). Malignant mucosa shows smudged intercellular spaces, atypical,

tortuous vessels, and fluorescein leakage (**c**). Kinetic Graph of a 60 frame sequence (**b**, **d**). The curves almost overlap and represent the homogeneous tissue typical of healthy epithelium (**b**), whereas the curves in frames of malignant mucosa differ significantly (**d**). ROI region of interest

cases, safe margin resection could be performed independent of the use of CLE. In-sano resection was defined by circular margin specimens in the intraoperative frozen section. A safety margin of >0.5 cm was confirmed in the definitive histopathological assessment in all patients.

Sequence selection, fluorescein flooding

With the Cellvizio Viewer 1.6.2, we assessed a kinetic graph of adequate Fluorescein flooding (Fig. 1). We recognized the initial Fluorescein enhancement after 208 s, and we obtained good quality images after 224 s. Out of 12,065 images, 60 sequences (3600 images; 33 sequences of SCC and 27

sequences of benign mucosa) were selected and evaluated as representative in acceptable quality. We matched each of these sequences (60 images, 5 s) with a corresponding sample by H&E staining to determine the diagnostic accuracy. All benign mucosal specimens were free of dysplasia or carcinoma in situ. The mean image acquisition time was 5 min (SD = 1.3) for each patient, supplemented by the assembly and disassembly of the scan unit of an additional approximately 5 min.

Tissue homogeneity in benign and malignant mucosa

Tissue homogeneity was defined as variations in the gray-scale values between the defined quadrangular ROIs. These values were obtained as a kinetic graph (Figs. 1 and 2) and a detailed numeric list. Benign and malign frames had a mean and SD of the mean of signal intensity or gray-values of 232.1 ± 3.34 and 467.3 ± 9.72 , respectively ($P < 0.001$). As a measure of variability between the four ROIs, the mean of SD in the whole 5 s/60 frame sequence was 39.1 ± 1.03 and 101.5 ± 2.6 for the healthy epithelium and SCC, respectively ($P < 0.001$).

The results of each rater are presented in Table 2. Inhomogeneous tissue architecture was observed on average in 14.8% (3.7–48.1%) of CLE sequences recorded in healthy mucosa and in 81.8% (45.4–96.9%) in CLE sequences recorded in SCC (chi-square test: $P < 0.001$). Overall sensitivity and specificity of 81.8% (95%CI: 76.3–86.2) and 86.2% (95%CI: 80.6–90.4) could be achieved, considering tissue homogeneity as the sole criterion. The sensitivity and specificity values varied from 59.1% to 96.9% and 51.8% to 100%, respectively (Table 2). The interrater reliability can be classified as moderate with a κ value of 0.53.

Discussion

This study evaluated the diagnostic value of tissue inhomogeneity of probe-based CLE in diagnosing laryngeal SCC. Carcinoma's important histological features generally include heterogeneous areas with necrosis, peritumoral inflammation, atypical mitosis, and epithelial beads [24]. Most proposed scoring systems for evaluating CLE images include tissue inhomogeneity as a conjectured correlation to the typical histological features [16–19]. An objective analysis of this feature has, however, to the best of our knowledge, not yet been performed. In the present study, we demonstrate a difference in tissue homogeneity given by the SD of the 4 ROIs in the field of view of the CLE frames with values of 39.1 ± 1.03 and 101.5 ± 2.6 for the healthy epithelium and SCC, respectively ($P < 0.001$). The identical sequences were also evaluated by seven blinded raters which identified, on

average, tissue inhomogeneity in 14.8% of CLE sequences recorded in healthy mucosa and 81.8% in SCC ($P < 0.001$) and is, therefore, in line with the objective method using the kinetic graph from the Cellvizio Software. With the help of this software, the optimal point of examination could also be assessed, as tissue flooding of fluorescein is noticeable after 208 s and peaks at 224 s. This information can help us to manage and improve the time-efficiency of intraoperative examination with CLE.

We also evaluated in this study tissue homogeneity as a stand-alone feature for the diagnosis of SCC and obtained by our seven blinded raters an overall sensitivity and specificity of 81.8% and 86.2% with a moderate interrater variability of 0.53.

We found tissue homogeneity alone to be susceptible to motion artifacts (Fig. 3 e, f). As depicted in the image, motion artifacts can cover the whole frame or some sections resulting in cells appearing stretched. This issue can easily be ignored by the clinicians performing CLE in the intended setting, i.e., real-time, in-vivo for evaluation during or before oncologic surgery, as the movement of the probe relative to the underlying tissue can easily be perceived. This study shows that an automatic examination with ROI is susceptible to these movements. These translate into a more considerable SD value between the different regions of interest (Fig. 3d–f). Although a manual selection of the frames can mitigate these problems, other solutions based on artificial intelligence and deep learning algorithms that are also in development by our group are more appropriate to solve this issue [25]. The approach by Aubreville et al., based on transfer learning from intermediate endpoints within a pre-trained Inception v3 network with tailored preprocessing and five-fold cross-validation, showed an overall accuracy of 94.8% in identifying image slices tainted by motion artifacts [25]. The automatic elimination of motion artifacts can improve the classification of images for diagnosis of SCC and will be a subject of future investigations. Other artifacts such as blood, saliva, air bubbles, and debris on the surface of the probe can also taint the quality of the images and simulate regional inhomogeneity (Fig. 3a–c). AI-based approaches to eliminating these relatively common artifacts are also in development and could improve classification accuracy.

The most distinctive histological characteristic of invasive carcinoma is the penetration of the basal membrane. Due to the fixed depth of 60 μm of this technic, the basal membrane cannot be correctly evaluated. Therefore, we cannot uncritically incorporate the interpretation of CLE images from the knowledge of histopathology, and we must verify the conjectured malignancy criteria. Abbaci et al. recently showed several examples of discrepancies between CLE interpretation and final histological diagnosis by a group of 3 pathologists, such as HNSCC being interpreted as respiratory epithelium and vice versa amounting to a sensitivity and specificity of,

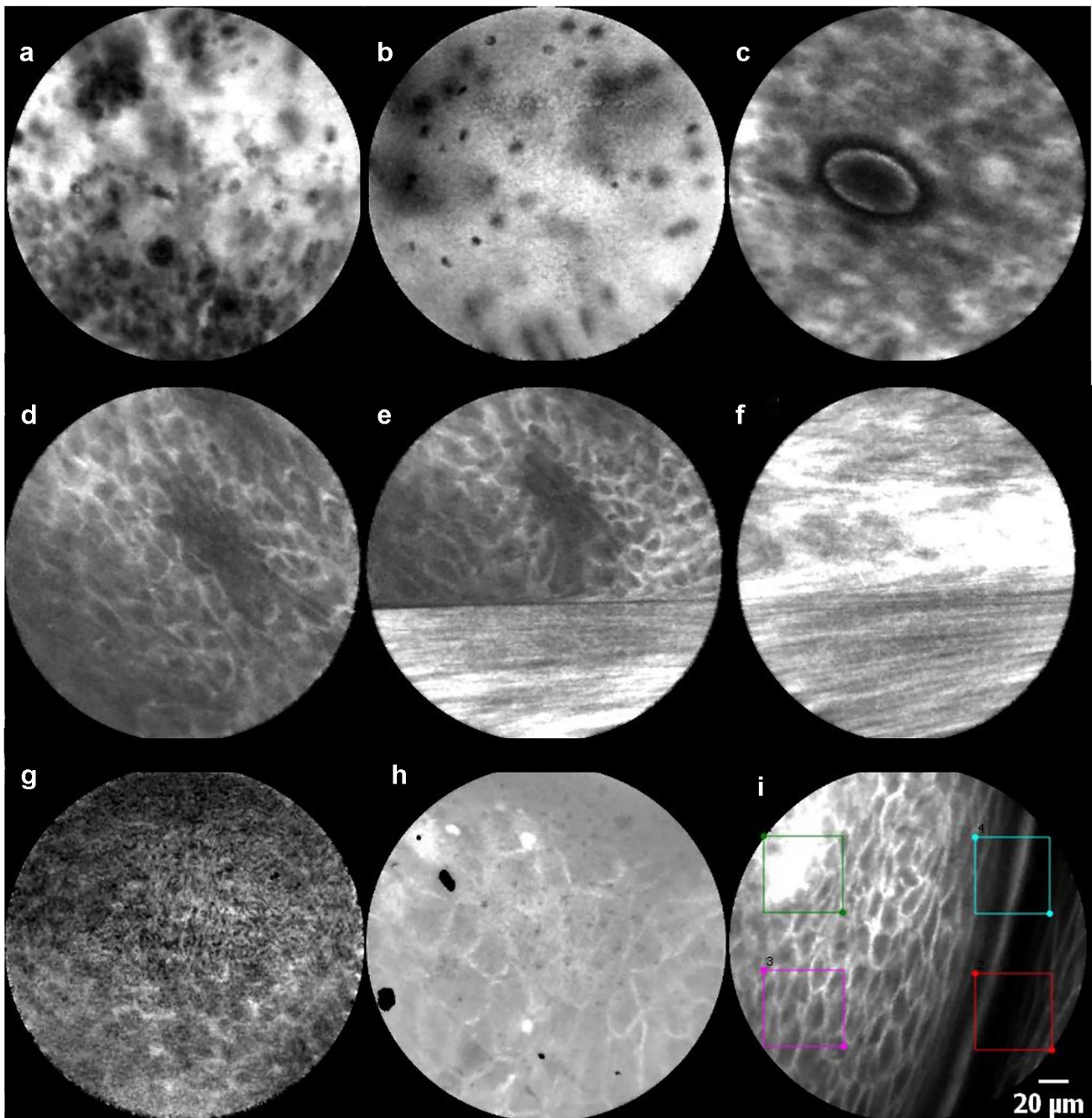


Fig. 3 CLE image artifacts with the potential to taint image quality and cause a higher variability measured by ROI. Contamination of the probe with blood (a) or saliva (b). An air bubble due to fluid between the probe and the mucosa. Motion artifacts due to slippage of the probe (d–e); normal mucosa with the classic honeycomb pat-

tern (d) and the two following frames with motion artifacts (e and f). Contact loss of the probe from the mucosa (g). Dirt on the lens complicates the visualization of the mucosa (h). Tangent position of the probe on the tissue (i) with an illustration of the four ROIs. The gray values of the individual ROIs differ in this frame

respectively, 73.2–75% and 30–57.4%, respectively [26]. A score developed by our group to larynx and pharynx, which took tissue homogeneity, cell size, presence of cell clusters, evaluation of capillary vessels, cell size, and cell borders into consideration, showed a sensitivity and specificity of 95.1% and 86.4% for clinicians with experience in the

technique, irrespective of background (pathologist or head and neck surgeon) and 86.4% and 86.1% for CLE examiners without previous experience with the method [17]. Another score developed by Oetter et al. specifically for the oral region, considering tissue homogeneity, intercellular gaps, cell morphology, fluorescence leakage, and vessel shape,

showed similar sensitivity and specificity of 95.3% and 88.9%, respectively [16]. Although these diagnostic metrics do not enable CLE to be a substitute for histopathology, if nothing else, because only a statement regarding presence or absence of malignancy is possible at the moment, they do come close to the values of intraoperative frozen sections, which show accuracy levels of 98.6–88.2% [27, 28].

In a previous study about the diagnostic value of intraepithelial capillary loops and atypical vessels in SCC diagnosis, we demonstrated that a cutoff value for vessel diameter of 30 μm enables the diagnosis of SCC with a sensitivity and specificity of, respectively, 90.6% and 71.3% [29]. However, while both criteria, tissue homogeneity and vessel diameter, as stand-alone diagnostic features, provide good sensitivity and specificity values, its diagnostic metrics are still inferior to diagnostic scores that integrate both features [17].

Histological examination is performed with a vertical slice through the different tissue layers. In contrast, probe-based CLE provides a tangential depiction of the epithelium at a fixed depth with a defined limited field of view [14, 17, 22]. With its 5–10 cell layers, the laryngeal epithelium has an average of 150 μm thickness [30, 31]. Arens et al. showed in a total of 206 vocal fold lesions that a progressive thickening of vocal cord epithelium over the different grades of dysplasia occurs with early invasive carcinoma showing an average epithelial thickness of 974 μm with a minimum of 800 μm [30]. However, the fixed depth of examination of 60 μm is not sufficient to evaluate the basal membrane. Therefore, this determinant in the histopathological differentiation between carcinoma in situ and invasive carcinoma cannot be used. Currently, there is no data comparing carcinoma in situ with invasive carcinoma using tissue homogeneity and capillary structure criteria [32]. However, this should be part of future research.

The rationale for non-invasive in vivo diagnostics is to define and control the resection margins precisely. The probe enables scanning the entire surgical field in just a few minutes. It allows the surgeon to get an overview of the extent of resection even with three-dimensional defects. The CLE may increase the probability of achieving an in-sano resection at the "first strike" and decrease the risk for re-resection, as the positive resection margin appeared to be an independent adverse prognostic factor [33]. In addition, the CLE can help take frozen sections in a more targeted manner to ensure the safety of the in-sano resection in communication with the pathologist. Thus, CLE offers the possibility to keep the resection volume as low as possible and, at the same time to realize a sufficient assessment of the entire resection margin. Furthermore, the clear differentiation between benign mucosal changes, normal mucosa, and malignant changes

is essential in early tumor diagnostic. In addition, CLE significantly helps to improve tumor follow-up monitoring for possible local recurrence and reduces the risk caused by unnecessary biopsies.

Other optical techniques, such as narrow-band imaging (NBI), have shown promising results in tumor diagnosis in the upper aerodigestive tract with an overall sensitivity and specificity of 89% and 96% [34]. This technique differs fundamentally from CLE. Due to its natural contrasted evaluation of vascular patterns in large mucosal areas [35] it can be considered a horizontal method to easily detect suspicious mucosal regions compared to white light endoscopy [36]. The magnification provided by NBI does not allow for the depiction of the cellular level. In addition to visualizing mucosal capillaries and atypical, aberrant vessels, CLE can also visualize superficial epithelial cells' contour and tissue architecture with up to $\times 1000$ magnification, resulting in an additional dimension for healthy mucosa and squamous cell carcinoma differentiation contrast to NBI. CLE, however, only enables the examination of a small section of a few millimeters at a time (2.6 mm diameter in this study) and requires the application of a fluorescent dye. CLE with simultaneous NBI-endoscopy could potentially enable a more accurate diagnostic. However, there are no studies at the moment that show the benefit of the combined application of both techniques.

Conclusions

The precise analysis of tissue homogeneity is a promising approach to improving CLE's diagnostic value to differentiate malignant squamous epithelial lesions and benign mucosa. Malignant lesions often have an inhomogeneous appearance on CLE examination and can be identified based on this feature with 81.8% and 86.2% sensitivity and specificity. Furthermore, the in-vivo, real-time examination by the clinician performing the technique during oncological surgery mitigates some weaknesses of the automatic analysis of tissue homogeneity in whole sequences, such as motion artifacts, saliva, blood, and debris, which can be directly identified as the cause for an apparent inhomogeneity. In addition, CLE enables the depiction of other criteria, which, when taken together in consideration, can provide a higher diagnostic value as a single-criterion examination considering only tissue homogeneity.

Acknowledgements This project was supported by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) grant number GO 3182/2-1, MA 4898/17-1, OE 743/1-1, STE 1877/7-1; Project Number 439264659.

Funding Open Access funding enabled and organized by Projekt DEAL.

Declarations

Conflict of interest None of the authors has a personal conflict of interest to declare.

Ethical approval All procedures performed in this study involving human participants were following the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Ethical approval (approval number 60_14 B) was given by the Non-Intervention Clinical Research Ethics Committee of the Medical Faculty of FAU Erlangen-Nürnberg.

Informed consent A formal informed consent procedure was obtained from all study participants.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Maier H, Dietz A, Gewelke U, Heller WD, Weidauer H (1992) Tobacco and alcohol and the risk of head and neck cancer. *Clin Investig* 70:320–327. <https://doi.org/10.1007/BF00184668>
2. De Leeuw F, Abbaci M, Casiraghi O, Ben Lakhdar A, Alfaro A, Breuskin I, Laplace-Builhé C (2000) Value of full-field optical coherence tomography imaging for the histological assessment of head and neck cancer. *Lasers Surg Med* 52:768–778. <https://doi.org/10.1002/lsm.23223>
3. Navarrete-Dechent C, Aleissa S, Cordova M, Liopyris K, Sahu A, Rossi AM, Lee EH, Nehal KS (2021) Management of complex head-and-neck basal cell carcinomas using a combined reflectance confocal microscopy/optical coherence tomography: a descriptive study. *Arch Dermatol Res* 313:193–200. <https://doi.org/10.1007/s00403-020-02037-6>
4. Sievert M, Oetter N, Auberville M, Stelzle F, Maier A, Eckstein M, Mantsopoulos K, Gostian AO, Mueller SK, Koch M, Agaimy A, Iro H, Goncalves M (2021) Feasibility of intraoperative assessment of safe surgical margins during laryngectomy with confocal laser endomicroscopy: a pilot study. *Auris Nasus Larynx* 48:764–769. <https://doi.org/10.1016/j.anl.2021.01.005>
5. Dittberner A, Ziadat R, Hoffmann F, Pertzborn D, Gassler N, Guntinas-Lichius O (2021) Fluorescein-guided panendoscopy for head and neck cancer using handheld probe-based confocal laser endomicroscopy: a pilot study. *Front Oncol* 11:671880. <https://doi.org/10.3389/fonc.2021.671880>
6. Neumann H, Kiesslich R, Wallace MB, Neurath MF (2010) Confocal laser endomicroscopy: technical advances and clinical applications. *Gastroenterology* 139:388–392. <https://doi.org/10.1053/j.gastro.2010.06.029>
7. Nabi Z, Reddy DN (2019) Optical biopsy in gastroenterology: focus on confocal laser endomicroscopy. *Indian J Gastroenterol* 38:281–286. <https://doi.org/10.1007/s12664-019-00986-z>
8. Xiong YQ, Ma SJ, Zhou JH, Zhong XS, Chen Q (2016) A meta-analysis of confocal laser endomicroscopy for the detection of neoplasia in patients with Barrett's esophagus. *J Gastroenterol Hepatol* 31:1102–1110
9. Tontini GE, Mudter J, Vieth M, Günther C, Milani V, Atreya R, Rath T, Nägel A, Hatem G, Sturmiolo GC, Vecchi M, Neurath MF, Galle PR, Buda A, Neumann H (2018) Prediction of clinical outcomes in crohn's disease by using confocal laser endomicroscopy: results from a prospective multicenter study. *Gastrointest Endosc* 87:1505–1514. <https://doi.org/10.1016/j.gie.2017.10.033>
10. Günther U, Daum S, Heller F, Schumann M, Loddenkemper C, Grünbaum M, Zeitz M, Bojarski C (2010) Diagnostic value of confocal endomicroscopy in celiac disease. *Endoscopy* 42:197–202
11. Vasilev I, Mamenko I, Tabanakova I, Vikulova I, Shevel V, Ushkov A, Novickaya T, Archakova L, Sokolovich E, Yablonskii P (2018) Probe-based confocal laser endomicroscopy in metastatic pulmonary calcification. *J Bronchol Interv Pulmonol* 25:60–62
12. Kriegmair MC, Ritter M, Michel MS, Bolenz C (2017) Moderne endoskopische Bildgebungsverfahren für das Urothelkarzinom der Harnblase [Modern endoscopic imaging tools for urothelial carcinoma of the urinary bladder]. *Aktuelle Urol* 4:296–305
13. Kamen A, Sun S, Wan S, Kluckner S, Chen T, Gigler AM, Simon E, Fleischer M, Javed M, Daali S, Igressa A, Charalampaki P (2016) Automatic differentiation based on confocal endomicroscopic images for intraoperative guidance in neurosurgery. *Biomed Res Int*. <https://doi.org/10.1155/2016/6183218>
14. Sievert M, Auberville M, Oetter N, Stelzle F, Maier A, Mantsopoulos K, Iro H, Goncalves M (2021) Confocal laser endomicroscopy of head and neck squamous cell carcinoma: a systematic review. *Laryngorhinotologie*. <https://doi.org/10.1055/a-1339-1635>
15. De Palma GD (2009) Confocal laser endomicroscopy in the “in vivo” histological diagnosis of the gastrointestinal tract. *World J Gastroenterol* 15:5770–5775. <https://doi.org/10.3748/wjg.15.5770>
16. Oetter N, Knipfer C, Rohde M, von Wilmowsky C, Maier A, Brunner K, Adler W, Neukam FW, Neumann H, Stelzle F (2016) Development and validation of a classification and scoring system for the diagnosis of oral squamous cell carcinomas through confocal laser endomicroscopy. *J Transl Med* 14:159. <https://doi.org/10.1186/s12967-016-0919-4>
17. Sievert M, Mantsopoulos K, Mueller SK, Eckstein M, Rupp R, Auberville M, Stelzle F, Oetter N, Maier A, Iro H, Goncalves M (2021) Systematic interpretation of confocal laser endomicroscopy: larynx and pharynx confocal imaging score. *Acta Otorhinolaryngol Ital*. <https://doi.org/10.14639/0392-100X-N1643>
18. Sievert M, Mantsopoulos K, Mueller SK, Rupp R, Eckstein M, Stelzle F, Oetter N, Maier A, Auberville M, Iro H, Goncalves M (2021) Validation of a classification and scoring system for the diagnosis of laryngeal and pharyngeal squamous cell carcinomas by confocal laser endomicroscopy. *Braz J Otorhinolaryngol*. <https://doi.org/10.1016/j.bjorl.2021.06.002>
19. Nathan CA, Kaskas NM, Ma X, Chaudhery S, Lian T, Moore-Medlin T, Shi R, Mehta V (2014) Confocal laser endomicroscopy in the detection of head and neck precancerous lesions. *Otolaryngol Head Neck Surg* 151:73–80. <https://doi.org/10.1177/0194599814528660>
20. Moore C, Mehta V, Ma X, Chaudhery S, Shi R, Moore-Medlin T, Lian T, Nathan CA (2016) Interobserver agreement of confocal laser endomicroscopy for detection of head and neck neoplasia. *Laryngoscope* 126:632–637. <https://doi.org/10.1002/lary.25646>
21. Goncalves M, Iro H, Dittberner A, Agaimy A, Bohr C (2017) Value of confocal laser endomicroscopy in the diagnosis of vocal cord lesions. *Eur Rev Med Pharmacol Sci* 21:3990–3997
22. Goncalves M, Auberville M, Mueller SK, Sievert M, Maier A, Iro H, Bohr C (2019) Probe-based confocal laser endomicroscopy in

- detecting malignant lesions of vocal folds. *Acta Otorhinolaryngol Ital* 39:389–395. <https://doi.org/10.14639/0392-100X-2121>
23. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33:159–174
 24. Ciolofan MS, Vlăescu AN, Mogoantă CA, Ioniță E, Ioniță I, Căpitănescu AN, Mitroi MR, Anghelina F (2017) Clinical, histological and immunohistochemical evaluation of larynx cancer. *Curr Health Sci J* 43:367–375. <https://doi.org/10.12865/CHSJ.43.04.14>
 25. Aubreville M, Stoeve M, Oetter N, Goncalves M, Knipfer C, Neumann H, Bohr C, Stelzle F, Maier A (2019) Deep learning-based detection of motion artifacts in probe-based confocal laser endomicroscopy images. *Int J Comput Assist Radiol Surg* 14:31–42. <https://doi.org/10.1007/s11548-018-1836-1>
 26. Abbaci M, Casiraghi O, Vergez S, Maillard A, Lakhdar AB, De Leeuw F, Crestani S, Ngo C, Koscielny S, Ferchiou M, Assouly N, Temam S, Laplace-Builhé C, Breuskin I (2021) Diagnostic accuracy of in vivo early tumor imaging from probe-based confocal laser endomicroscopy versus histologic examination in head and neck squamous cell carcinoma. *Clin Oral Investig*. <https://doi.org/10.1007/s00784-021-04156-4>
 27. Layfield EM, Schmidt RL, Esebua M, Layfield LJ (2018) Frozen section evaluation of margin status in primary squamous cell carcinomas of the head and neck: a correlation study of frozen section and final diagnoses. *Head Neck Pathol* 12:175–180. <https://doi.org/10.1007/s12105-017-0846-6>
 28. Kubik MW, Sridharan S, Varvares MA, Zandberg DP, Skinner HD, Seethala RR, Chiosea SI (2020) Intraoperative margin assessment in head and neck cancer: a case of misuse and abuse? *Head Neck Pathol* 14:291–302. <https://doi.org/10.1007/s12105-019-01121-2>
 29. Sievert M, Eckstein M, Mantsopoulos K, Mueller SK, Stelzle F, Aubreville M, Oetter N, Maier A, Iro H, Goncalves M (2021) Impact of intraepithelial capillary loops and atypical vessels in confocal laser endomicroscopy for the diagnosis of laryngeal and hypopharyngeal squamous cell carcinoma. *Eur Arch Otorhinolaryngol*. <https://doi.org/10.1007/s00405-021-06954-8>
 30. Arens C, Glanz H, Wönckhaus J, Hersemeyer K, Kraft M (2007) Histologic assessment of epithelial thickness in early laryngeal cancer or precursor lesions and its impact on endoscopic imaging. *Eur Arch Otorhinolaryngol* 264:645–269. <https://doi.org/10.1007/s00405-007-0246-8>
 31. Armstrong WB, Ridgway JM, Vokes DE, Guo S, Perez J, Jackson RP, Gu M, Su J, Crumley RL, Shibuya TY, Mahmood U, Chen Z, Wong BJ (2006) Optical coherence tomography of laryngeal cancer. *Laryngoscope* 116:1107–1113. <https://doi.org/10.1097/01.mlg.0000217539.27432.5a>
 32. Sievert M, Stelzle F, Aubreville M, Mueller SK, Eckstein M, Oetter N, Maier A, Mantsopoulos K, Iro H, Goncalves M (2021) Intraoperative free margins assessment of oropharyngeal squamous cell carcinoma with confocal laser endomicroscopy: a pilot study. *Eur Arch Otorhinolaryngol* 278:4433–4439. <https://doi.org/10.1007/s00405-021-06659-y>
 33. Szewczyk M, Golusinski W, Pazdrowski J, Masternak M, Sharma N, Golusinski P (2017) Positive fresh frozen section margins as an adverse independent prognostic factor for local recurrence in oral cancer patients. *Laryngoscope* 128:1093–1098. <https://doi.org/10.1002/lary.26890>
 34. van Schaik JE, Halmos GB, Witjes MJH, Plaat BEC (2021) An overview of the current clinical status of optical imaging in head and neck cancer with a focus on narrow band imaging and fluorescence optical imaging. *Oral Oncol*. <https://doi.org/10.1016/j.oraloncology.2021.105504>
 35. Chiam KH, Shin SH, Choi KC, Leiria F, Militz M, Singh R (2021) Current status of mucosal imaging with narrow-band imaging in the esophagus. *Gut and liver* 15:492–499. <https://doi.org/10.5009/gnl20031>
 36. Betz CS, Kraft M, Arens C, Schuster M, Pfeffer C, Rühm A, Stepp H, Englhard A, Voggler V (2016) Optische diagnoseverfahren zur Tumorfrühdagnostik im oberen Luft-Speise-Weg. *HNO* 64:41–48

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.