

# Journal of Biomedical Optics

BiomedicalOptics.SPIEDigitalLibrary.org

## Dual-mode endomicroscopy for detection of epithelial dysplasia in the mouth: a descriptive pilot study

Nico Bodenschatz  
Catherine F. Poh  
Sylvia Lam  
Pierre Lane  
Martial Guillaud  
Calum E. MacAulay

**SPIE.**

Nico Bodenschatz, Catherine F. Poh, Sylvia Lam, Pierre Lane, Martial Guillaud, Calum E. MacAulay, "Dual-mode endomicroscopy for detection of epithelial dysplasia in the mouth: a descriptive pilot study," *J. Biomed. Opt.* **22**(8), 086005 (2017), doi: 10.1117/1.JBO.22.8.086005.

# Dual-mode endomicroscopy for detection of epithelial dysplasia in the mouth: a descriptive pilot study

Nico Bodenschatz,<sup>a</sup> Catherine F. Poh,<sup>a,b</sup> Sylvia Lam,<sup>a</sup> Pierre Lane,<sup>a</sup> Martial Guillaud,<sup>a</sup> and Calum E. MacAulay<sup>a,\*</sup>

<sup>a</sup>British Columbia Cancer Research Centre, Department of Integrative Oncology, Vancouver, British Columbia, Canada

<sup>b</sup>The University of British Columbia, Faculty and Dentistry, Vancouver, British Columbia, Canada

**Abstract.** Dual-mode endomicroscopy is a diagnostic tool for early cancer detection. It combines the high-resolution nuclear tissue contrast of fluorescence endomicroscopy with quantified depth-dependent epithelial backscattering as obtained by diffuse optical microscopy. In an *in vivo* pilot imaging study of 27 oral lesions from 21 patients, we demonstrate the complementary diagnostic value of both modalities and show correlations between grade of epithelial dysplasia and relative depth-dependent shifts in light backscattering. When combined, the two modalities provide diagnostic sensitivity to both moderate and severe epithelial dysplasia *in vivo*. © 2018 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JBO.22.8.086005]

Keywords: endomicroscopy; biomedical optics; fluorescence; light scattering; *in vivo* pathology.

Paper 170247R received Apr. 17, 2017; accepted for publication Jul. 25, 2017; published online Aug. 19, 2017.

## 1 Introduction

It is known that during early precancerous development in the squamous epithelium, structural phenotypic changes are often initially observable in cells and cell nuclei adjacent to the epithelial basement membrane.<sup>1,2</sup> These alterations are frequently denoted as “mild” dysplasia in their earliest stage. If mild dysplasia transforms further, abnormal cells and cell packing alterations become more prominent also in the middle and upper parts of the epithelium and are termed “moderate” and “severe” dysplasia, respectively. Throughout this precancerous progression from mild to moderate to severe degree, the altered tissue is increasingly likely to develop into carcinoma *in situ* (CIS) and eventually into invasive cancer.<sup>3</sup>

While precancerous tissue alterations may be macroscopically identifiable as areas of tissue abnormality by an experienced clinician, histopathological analysis of biopsies is usually necessary to distinguish lesions with high or low risk of developing into cancer. Our dual-mode endomicroscopy (DME) imaging modality aims at detecting and delineating tissue abnormality while providing real-time feedback on when and where to excise tissue for histopathological analysis. At the same time, we seek to address the clinical need for the detection of moderate and severe dysplasia to prevent their potential growth into invasive cancer.

The imaging modality described in this study combines fluorescence endomicroscopy (FE)<sup>4</sup> with a recently pioneered modality called diffuse optical microscopy (DOM).<sup>5</sup> The rationale for combining these two imaging modalities in a unified system lies in their complementary sensitivity to the structural changes associated with precancerous development. FE can image nuclear structure of the most superficial epithelial cell layers by using a nuclei staining extrinsic fluorophore dye.<sup>6–10</sup> This dye contrast helps to reveal severe dysplasia through the analysis of nuclear size and density in the surface layers of the epithelium.<sup>11–14</sup> However, FE is mostly unable to detect

moderate dysplasia where structural alterations are limited to deeper cell layers. Therefore, we combine FE with DOM, which images and quantifies light backscattering from different tissue depths through a structured illumination approach.<sup>5</sup> Selection of structured illumination patterns allows for tailoring light penetration to diagnostically meaningful depths within the epithelium for added sensitivity toward moderate dysplasia.

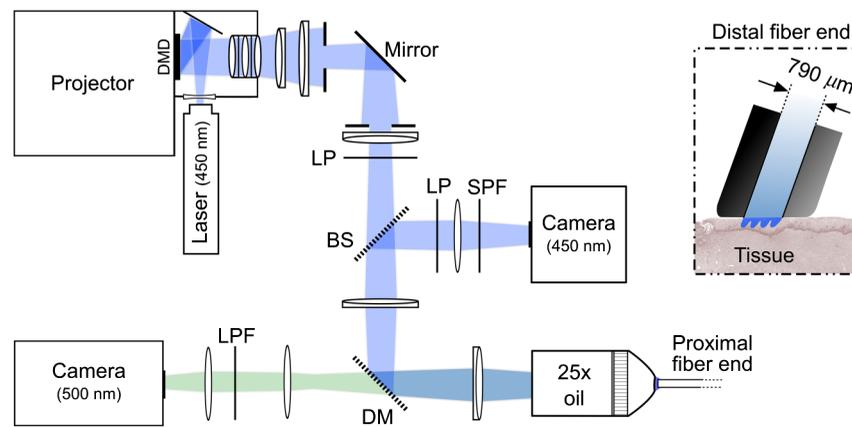
In the following sections, we introduce the operational principles of our new imaging modality and provide details of the *in vivo* pilot study (Sec. 2). In Sec. 3, we demonstrate the ability of DME to detect structural changes and to associate them with grade of dysplasia of lesions in the oral cavity. The presented qualitative and quantitative imaging results are in agreement with our recent findings in the cervix, where we used DOM and FE independently.<sup>5,11</sup> Sections 4 and 5 discuss the significance and limits of our work and conclude on our results, respectively.

## 2 Materials and Methods

### 2.1 Device and Imaging Approach

The schematic setup of our optical DME is given in Fig. 1. A projection unit (DLP LightCrafter, Young Optics Inc., Taiwan) featuring a digital micromirror device (DMD) is used along with a blue laser diode module (Flexpoint, Laser Components, Germany) to create sinusoidal projection patterns. To achieve wide field projection, the point laser beam is widened and directed to the DMD using an optical diffuser plate and a lens. Numerous lenses are in place to demagnify the structured illumination patterns into a 25× oil immersion microscope objective (Zeiss Plan-Neofluar, Germany). Tissue light projection is achieved with a coherent fiber bundle (FIGH-30-850N, Myriad Fiber Imaging) comprising 30,000 imaging fibers with a circular imaging field of view of 790 μm in diameter. The polished proximal end of this 1.5-m long bundle is placed at the working distance of the objective and index-matching

\*Address all correspondence to: Calum E. MacAulay, E-mail: [cmacula@bccrc.ca](mailto:cmacula@bccrc.ca)



**Fig. 1** Schematic diagram of the DME system employing two digital cameras, a laser diode module, and a projection system. A DM and a BS achieve separation of fluorescence and backscattering. LP, an SPF, and an LPF help to purify signals.

immersion oil with a refractive index of  $n = 1.51$  is placed between objective and fiber to suppress reflections at the fiber interface.<sup>15</sup> The distal end of the bundle is placed in contact with the tissue surface (see inset of Fig. 1), which allows for the light patterns to be projected onto the tissue. To avoid terminal reflections at the imaging fibers the distal end of the bundle is polished at an angle of 20 deg. We determined the fiber projection and detection numerical aperture to be  $0.27 \pm 0.02$ .

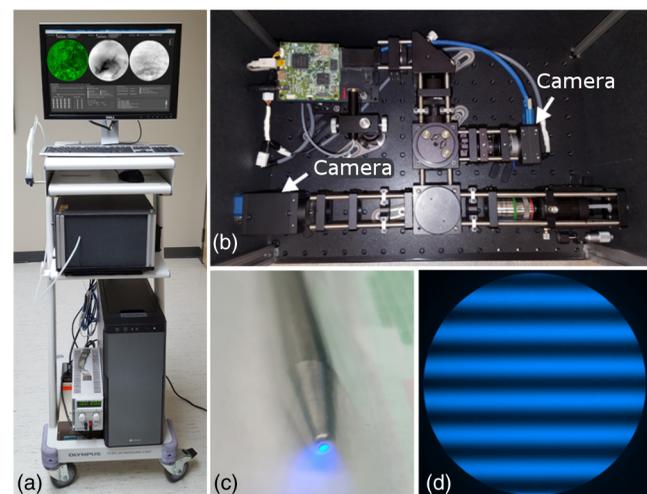
The blue excitation light encounters three major interaction processes relevant to our imaging device: fluorescence, backscattering, and absorption. A small portion of the light excites acriflavine hydrochloride (AH) fluorophores, which are topically applied to the tissue surface. Similar to many previous studies, we use an aqueous 0.01% solution of AH (Sigma-Aldrich).<sup>4,6,7,12,13</sup> This antiseptic dye temporarily binds nucleic acids thus providing contrast for the shape, size, and density of cell nuclei. The limited permeation of the dye into the surface 2 to 3 cell layers (i.e.,  $\approx 30 \mu\text{m}$ ) limits FE contrast to the tissue surface.<sup>11</sup> While very strong and prolonged illumination of stained cells may cause cellular photo damage,<sup>16</sup> AH has a reported long history of safe clinical usage as an antiseptic compound. AH molecules fluoresce with a spectrum centered at about  $\lambda = 500 \text{ nm}$ . This green fluorescence light is in part recaptured by the projection image bundle. In the optical setup, these wavelengths above 460 nm are transmitted through a dichroic mirror (DM, BrightLine Di02-R442, Semrock) and focused onto a digital camera (GS3-U3-51S5M, Point Grey, Canada;  $2448 \times 2048$  pixel resolution). A long pass filter (LPF) further purifies the captured fluorescence signal.

A second portion of the light is backscattered while retaining its wavelength. This reflectance signal comprises both diffusely and subdiffusely backscattered light.<sup>17</sup> If light gets scattered back into the image bundle and is thus directed back into the optical setup, it is partly guided to a second digital camera (CM3-U3-50S5M, Point Grey, Canada;  $2448 \times 2048$  pixel resolution) using a 50:50 beam splitter (BS). A short pass filter (SPF) prevents any residual fluorescence from being captured by this camera. To significantly reduce the extent by which specular reflections in the optical setup contribute to the reflectance signal, two crossed linear polarizers (LP) are built into the system. Note, however, that the fiber bundle is not polarization maintaining which makes the device insensitive to polarization-dependent light propagation effects.

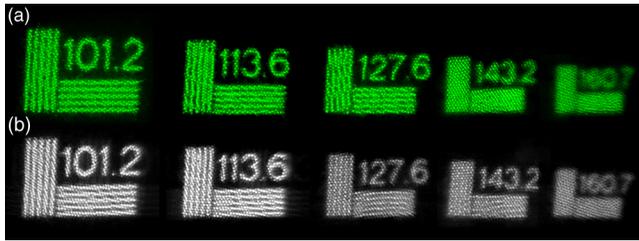
One significant technical merit of our DME approach is the possibility to implement both modalities in a single optical setup using a single light source and projection system. We make use of a laser diode module with an output power of 70 mW at  $\lambda = 445 \pm 5 \text{ nm}$ . The reason for the selected wavelength is given by the excitation spectrum of the acriflavine fluorophore dye that is used for FE. At the same time, this wavelength has a lower diffuse background signal in DOM as compared to red or infrared radiation and blood absorption effects give rise to vascular contrast.

Light absorption by intrinsic tissue molecules is the third interaction process that affects DOM, with hemoglobin being the dominant absorber at the 445-nm excitation wavelength. We find local dominance of these absorption effects to be restricted to vascularized areas and depths and thus observe vascular patterns in DOM images.

Photographs of the clinical DME system and its components are given in Fig. 2. Figure 2(a) presents the entire clinical system



**Fig. 2** (a) Photograph of the entire clinical DME system including computer (bottom), optical setup (middle), and graphical user interface (top). (b) A photograph of the optical setup reveals components as presented in Fig. 1. The imaging wand (c) carries the fiber bundle for light projection and detection. A close up picture (d) reveals a structured light pattern emanating from the tip of the probe.



**Fig. 3** (a) Composed image showing the maximum resolution attainable in the fluorescence mode and (b) in the reflectance mode as measured by white light trans-illumination of a resolution target. Numbers correspond to line pairs per millimeter.

and Fig. 2(b) is a photograph of the optical setup that corresponds to the schematic diagram in Fig. 1. We use a 17-cm long stainless steel wand [Fig. 2(c)] for easy handling and disinfection. The fiber bundle is glued inside the wand and a close up of the probe tip reveals one-structured illumination pattern emanating from the circular imaging field of view [Fig. 2(d)].

Figure 3 demonstrates the principal lateral resolution for both imaging modalities as limited by the interfiber spacing. Using transillumination of a resolution target positioned at the distal fiber end, the finest line structure resolved for both the fluorescence detection arm (top row) and the reflectance detection arm (bottom row) is 127.6 line pairs/mm. This corresponds to a lateral resolution limit of  $7.8 \mu\text{m}$  (Rayleigh criterion), which is consistent with the center-to-center fiber spacing of  $4.3 \mu\text{m}$ .

As reflectance contrast in DOM is by design simultaneously obtained from multiple cell layers, the actual lateral and axial resolution are compromised and dependent on the spatial frequency of the projection pattern. Figure 4 approximates the associated light penetration depths for all spatial frequencies employed in our experiments. These sketched penetration depths are derived from Monte Carlo simulations and correspond to the depth not surpassed by 95% of photons.<sup>5</sup> It should be noted that the average penetration maximum of all photons for the given frequencies is only about half the indicated 95% penetration depth.

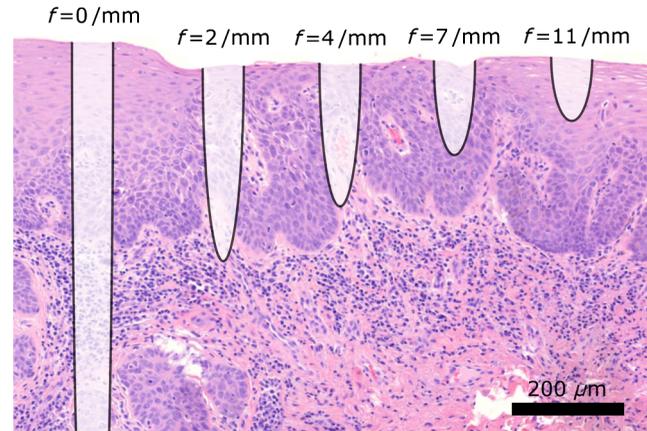
We employ structured illumination patterns of the form  $I(x, y) = \hat{I} \sin(2\pi f x + \varphi)$ , with 3-phase projections  $I_1, I_2, I_3$  per spatial frequency  $f$  using phase shifts  $\varphi$  of 0 deg, 120 deg, and 240 deg, respectively. For every phase projection triplet, we capture three reflectance and three fluorescence phase images in parallel. We denote the three reflectance phase images by  $X_1, X_2, X_3$ . Reflectance values  $R$  for a given spatial frequency and pixel location are then derived using

$$R(f) = \frac{\sqrt{2}}{3} \sqrt{(X_1 - X_2)^2 + (X_2 - X_3)^2 + (X_3 - X_1)^2}. \quad (1)$$

Reflectance at  $f = 0$  can be computed with any set of phase captures using

$$R_{f=0} = (X_1 + X_2 + X_3)/3. \quad (2)$$

As demonstrated in Fig. 4, reflectance at spatial frequency  $f = 0$  is sensitive to backscattering from beyond the epithelial thickness. The actual depth sensitivity at  $R_{f=0}$  is strongly dependent on the fiber bundle's imaging diameter. Because of its strong signal and sometimes interesting vascular absorption contrast patterns from larger tissue depths, we use  $R_{f=0}$  for qualitative image analysis.



**Fig. 4** Estimated depth of light penetration for tissue reflectance at various spatial frequencies in DOM. A sample tissue section of the mucosa from the ventral tongue illustrates how different spatial frequencies collect backscattering from different epithelial depths.

Recording a higher number of spatial frequencies reduces the image acquisition speed while providing more information on depth-dependent epithelial backscattering. In this pilot study, we selected four spatial frequencies [ $f = (2, 4, 7, 11) \text{ mm}^{-1}$ ], which require a total of 12-phase images (three images per frequency) to be captured in parallel by each camera for a comprehensive evaluation of one measurement site. Our multithreaded software and the two high-speed USB three cameras allow for a parallel image acquisition rate of 49 frames per camera and per second. As a consequence, DME continuous-mode imaging with the two forms of contrast covering the same imaging field of view is possible at a speed of 4.1 frames/s. This speed can be significantly increased by reducing the selection of spatial frequencies interrogated.

It was demonstrated by Keahey et al.<sup>18,19</sup> that structured illumination can similarly be used in FE to improve surface sensitivity and imaging contrast. In our case, however, the fluorescence signal mostly originates from very superficial layers of the stratified squamous epithelium of the mucosal surface and cannot be improved much by structured illumination as described above. We also observe imaging noise to be enhanced by Eq. (1) owing to the low optical output power at our probe tip of only  $200 \mu\text{W}$ . As a consequence, we use the fluorescence analog of Eq. (2) at the lowest spatial frequency for the computation of FE images.

## 2.2 Image Processing

Using our in-house C++ DME system software, numerous image postprocessing steps are implemented and carried out in real time to reduce image artifacts and to enhance perceived image quality. Images of both cameras are cropped to a  $2048 \times 2048$  pixel region of interest and are subject to  $2 \times 2$  hardware binning before data transmission. To further reduce the data size and in accordance with the different resolution constraints, we perform  $2 \times 2$  software binning on the fluorescence phase images and  $4 \times 4$  software binning on the reflectance phase captures and thus obtain images of  $512 \times 512$  and  $256 \times 256$ , respectively. As the effective resolution of both cameras is much higher than that of the 30,000 fiber imaging bundle, this size reduction enhances computational speed and takes away resolution not supported by the fiber bundle or imaging mode.

After image acquisition, we correct all phase images by subtraction of reference frames that were obtained by immersing the imaging wand in a dark water-filled container while running the structured illumination patterns. This dark reference allows us to quantify all internal reflections and fluorescence from within the optical system, including spurious light fluorescence and backscattering from within the fiber bundle and unwanted reflections occurring on every optical surface in the optical setup.

Owing to the limited image bundle diameter and the correspondingly small imaging field of view, the integral projection intensity for the different phase projections is slightly different. This gives rise to ringing artifacts in the computed reflectance image especially for low spatial frequencies. To prevent this artifact, we perform high-pass filtering on all phase images before using them in Eq. (1) as previously described in Refs. 5 and 20.

Not all areas within the circular imaging field of view encounter the same projection intensity and projection focus. Therefore,  $R(f)$  values are divided by flat fields subsequent to their computation with Eqs. (1) or (2). These flat fields are recordings of  $R(f)$  for freshly prepared intralipid emulsions with a reduced scattering coefficient of  $\mu'_s \approx 2.0 \text{ mm}^{-1}$  (3.0 g of Intralipid®-20% in 37.4 g of distilled water), which we acquired once at the beginning of our imaging study. As all the data were divided by identical flat fields and no system changes were made during the course of this study, this correction has no influence on our quantitative evaluations.

Fiber bending and tissue–probe interface variations influence projection and collection efficiencies of individual fibers and give rise to superimposed image noise patterns. As these noise patterns are often very similar in adjacent imaging iterations, we were able to enhance the perceived image quality of both fluorescence and reflectance images by subtracting a time lapse average of preceding imaging iterations. We used this real-time processing option of our imaging software only on displayed images and in our qualitative image analysis, but not on the data we used for quantitative evaluation.

### 2.3 Patients and Imaging

In the course of this pilot clinical study from January to March 2017, we have recruited 21 patients who also participated in either a pan Canadian study or a longitudinal study for oral pre-malignant lesions in the Oral Oncology Clinic of the BC Cancer Agency in Vancouver, Canada. This study was approved by the Ethics Board of the British Columbia Cancer Agency under the ethics number H11-00011. Both male (52%) and female patients (48%) between 58 and 84 years with an age average of 70 years were recruited on a voluntary basis. A total of eight biopsies were performed on the 21 patients. Imaging data of oral lesions with matched pathology-based diagnosis are therefore limited to these cases. Table 1 provides a summary of patient characteristics and biopsy diagnoses.

After obtaining a patient's informed consent, lesions of interest and a contralateral oral site were topically stained with AH using a cotton swab for  $\sim 10$  s. Subsequent to a mouth rinse with water, image recording started about 2 min after staining. Image acquisition took 5 to 10 s per imaging site and lasted 3 to 5 min in total depending on the number of sites. Between four and nine sites were imaged in each patient. Every site was measured by holding the imaging wand stable for at least 1 s in continuous recording mode. For every patient, an adjacent normal reference and in some cases, also a contralateral normal reference was recorded.

**Table 1** Overview of patient characteristics and biopsy diagnoses for lesions imaged in our pilot *in vivo* study.

Patients	21
Age	58 to 84 (mean 70)
Lesions imaged	27
Lesions with biopsy	8
Biopsy diagnoses	
Mild dysplasia	2
Moderate dysplasia	2
Severe dysplasia/CIS	1
SCC	2
Confounder (lichenoid mucositis)	1

Among the lesion sites, 74% were captured from the known high-risk anatomical sites (60 sites), including lateral tongue, ventral tongue, and floor of mouth. These areas are found to be among the most common sites of origin of squamous cell carcinoma (SCC) in the oral cavity.<sup>21</sup> Further imaging sites of this study include the buccal mucosa (12 sites), the palate (5 sites), and the gingiva (4 sites).

Subsequent to patient image acquisition, an intralipid emulsion (same emulsion as above) was imaged to confirm the consistency of the instrument response function and to keep track of laser intensity fluctuations.<sup>22</sup>

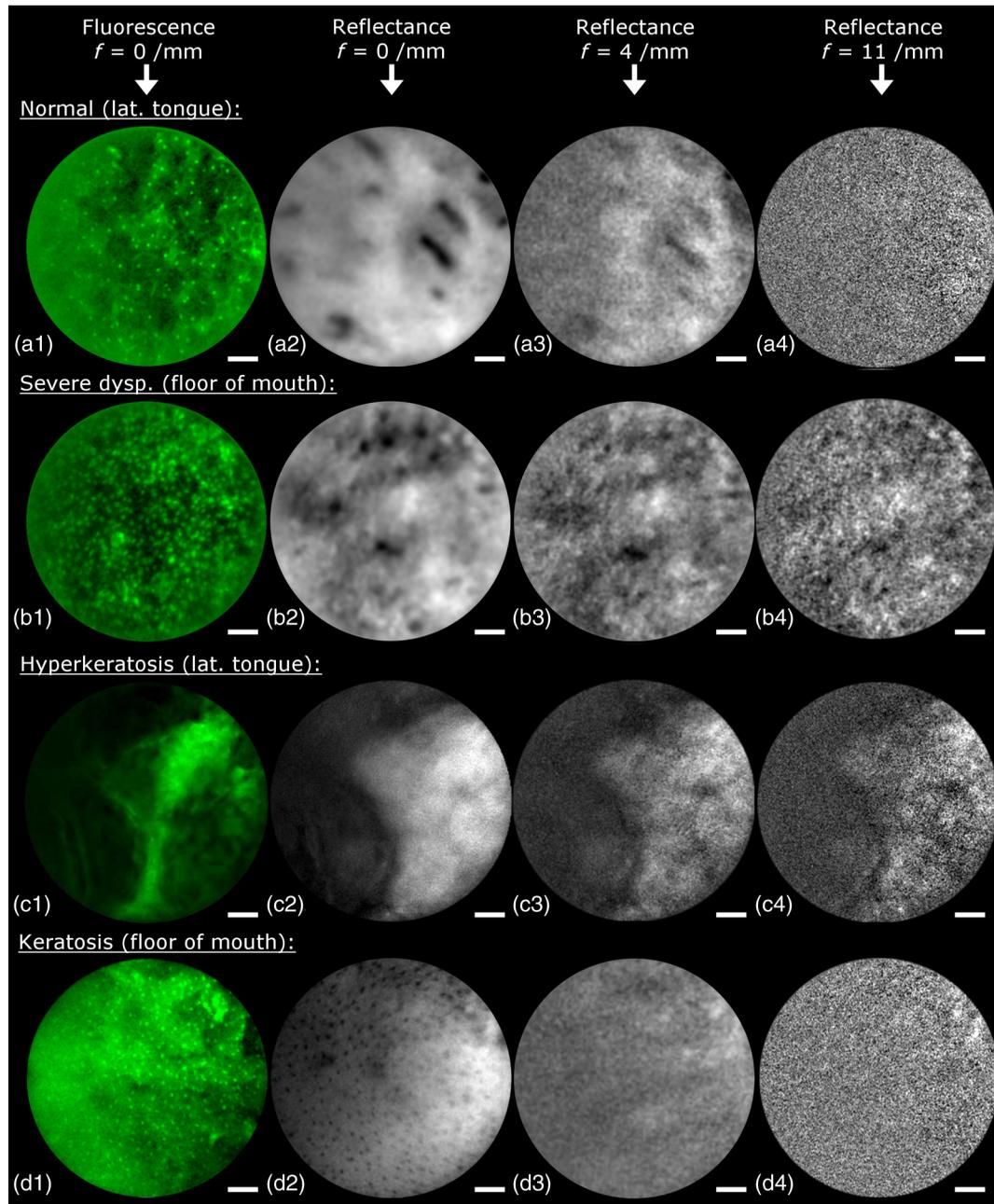
## 3 Results

### 3.1 Qualitative Analysis of Dual-Mode Endomicroscopy Images

In Fig. 5, we illustrate qualitative image contrast for four sets of DME images (a–d), which were captured in different patients and at different anatomical sites. For every site, we display one FE image (first column) and three DOM images (columns 2 to 4). Owing to the higher signal intensity and the improved signal-to-noise ratio, we generally find the contrast of DOM images to be best at the lowest spatial frequency (second column of Fig. 5). Inversely, reflectance at the spatial frequencies  $f = 7$  and  $11 \text{ mm}^{-1}$  is linked with lower signal intensities and very superficial depth sensitivity. This gives rise to a comparatively low signal-to-noise ratio. Especially, images a4 and d4 in Fig. 5 seem not to carry valuable structural information, yet their mean pixel intensities provide meaningful quantitative data in later analysis (see Sec. 3.2).

All presented images in Fig. 5 were captured in continuous acquisition mode while slightly moving the wand over the tissue. Cell nuclei appear as green dots in FE images and the small nuclear size and wide nuclear spacing [Fig. 5(a1)] are common features of a normal epithelium.<sup>23</sup> In all DOM images, bright areas correspond to increased backscattering, whereas darker areas are attributable to either less backscattering or vascular absorption of the illumination light.

A severely dysplastic epithelium [Fig. 5(b)] is known to feature nuclear crowding at the epithelial surface which can



**Fig. 5** Four sets of DME images (rows a–d) for different tissue sites comprising a FE image (first column) and three DOM images (columns 2–4) each. The three DOM images exemplify reflectance at frequencies  $f = 0, 4, \text{ and } 11 \text{ mm}^{-1}$ . Wide nuclear spacing (a1) and more homogenous backscattering (a2, a3, and a4) are features of normal epithelium. Nuclear crowding (b1) and heterogeneity in DOM (b2, b3, and b4) are features of severe dysplasia. Dark lines and dots (mostly in a2 and b2) are caused by blood absorption and represent blood vessels growing toward the epithelium from below. By contrast, vascular patterns are barely visible at high spatial frequency reflectance (columns 3 and 4) corresponding to the intended sampling of more superficial epithelial backscattering. Image series (c) and (d) present data from hyperkeratinized and keratinized epithelium, respectively. Scale bars measure  $100 \mu\text{m}$ .

be detected with FE [Fig. 5(b1)]. In addition, we observe DOM reflectance images to appear more heterogeneous as compared to corresponding images in a normal epithelium (compare b3 and b4 with a3 and a4 in Fig. 5). These intensity variations within DOM images of a severely dysplastic epithelium are likely to correspond to enlarged rete pegs. Such epithelial extensions are commonly found in high grade (i.e., moderate or severe) dysplasia and have also been reported in the form of

elevated standard deviation values in lateral optical coherence tomography (OCT) scans.<sup>24</sup> For all high-grade lesions imaged in our pilot study, we observe such increased lateral heterogeneity in the reflectance images.

In slightly more than 50% of all measured lesion sites, nuclei were not visible in FE such that nuclear size and spacing could not be evaluated [see example in Fig. 5(c1), a hyperkeratotic area of a lesion]. We attribute this lack in nuclear contrast to

enhanced tissue fluorescence from surface keratin. Both keratin autofluorescence and keratin staining by AH may be the cause of the overall increase in fluorescence.<sup>25</sup> We find that image-averaged fluorescence intensities in such areas of low nuclear visibility [Fig. 5(c1)] are often 5 to 10 times higher as compared to a normal epithelium with good nuclear visibility [Fig. 5(a1)]. Note that images in Fig. 5 have different intensity scaling. Interestingly, keratotic areas that appear bright in FE [Fig. 5(c1)] are similarly found to be bright in DOM [Figs. 5(c2) and (c3)]. Keratin backscattering seems to thus exceed the anticipated drop in reflectance that relates to light absorption by the stain AH.

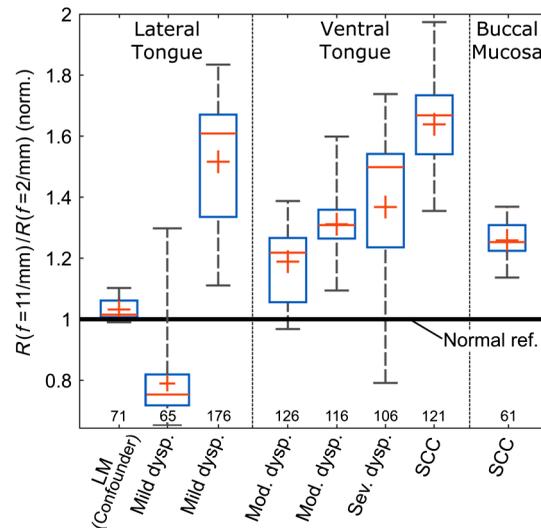
Similar to Fig. 5(c1), also Fig. 5(d1) features high fluorescence intensities. In this case, nuclei are visible in both FE [Fig. 5(d1)] and DOM [Fig. 5(d2)], which is a phenomenon that we repeatedly observed in moderately keratinized areas in patients and in healthy volunteers. Our hypothesis for the nuclear contrast in DOM is a strongly elevated scattering of the excitation light from cytoplasmic keratin, combined with light absorption by AH in cell nuclei as well as comparatively less scattering from cell nuclei.

### 3.2 Quantitative Dual-Mode Endomicroscopy Data Analysis

In a previous DOM study of the uterine cervix, we found a strong variability in absolute epithelial backscattering between patients and even in the same patient at neighboring tissue sites.<sup>5</sup> We observe the same phenomenon in the oral cavity, which likely relates to the known site-dependent variation in epithelial thickness.<sup>26,27</sup> To overcome this variability and to reduce the sensitivity to possible laser power fluctuations, we compute ratios of reflectance at different spatial frequencies. This is achieved by dividing reflectance images acquired with different spatial frequencies to obtain new images that encode scattering change with depth. It has been reported that changes in depth-dependent epithelial backscattering correlate with dysplastic tissue alterations.<sup>24,28</sup> Multiple reflectance ratios can be explored by different combinations of the four measured spatial frequency values. The ideal spatial frequency ratio is likely a function of epithelial thickness and thus of oral tissue site. For our imaging data, we observe highest sensitivity to epithelial dysplasia when using the ratio  $R(f = 11 \text{ mm}^{-1})/R(f = 2 \text{ mm}^{-1})$ . In a first approximation, this ratio relates backscattering from the upper epithelial cell layers to that from the entire epithelium. However, similar sensitivity can be obtained for many other frequency combinations except those from  $R(f = 0 \text{ mm}^{-1})$  which quantifies scattering from below the epithelium.

In Fig. 6, we evaluate our imaging data from all tissue biopsy sites corresponding to eight imaging sites in eight patients. For each of these sites, 61 to 176 frames were recorded in continuous imaging mode. We define a “frame” to be an independent measurement at a unique site where a fluorescence image and all reflectance spatial frequency images were acquired. Frames with motion artifacts in any one of their frame images were excluded from the analysis. Such artifacts were visually detected by scanning frame images for oscillatory intensity fluctuations. These ringing features develop if the corresponding 3-phase images are not acquired at the same tissue location.

For every frame, the reflectance ratio  $R(f = 11 \text{ mm}^{-1})/R(f = 2 \text{ mm}^{-1})$  was computed at every pixel and subsequently averaged over the central 80% of the imaging field of view. The latter constraint is meant to prevent projection boundary artifacts from influencing the data evaluation.<sup>5</sup> Image averaging



**Fig. 6** Relative change in depth-dependent epithelial backscattering with respect to a patient's adjacent or contralateral normal reference for all imaging sites with biopsy-confirmed diagnosis. Every column evaluates the backscattering ratio over the provided number of corresponding image frames recorded at one of eight biopsy sites. Mean values (red crosses), median values (red lines), 25% and 75% quartiles (blue boxes) and both minimum and maximum values (whiskers) are presented.

is motivated by the low signal-to-noise ratio of high spatial frequency reflectance images. Every box in the box-plot of Fig. 6 presents these image-averaged values for the frame ensemble of one measurement site. Average values (red crosses), median values (red lines), 25% and 75% quartiles (box ends), as well as minimum and maximum values (whiskers) are presented. To overcome the natural variability in epithelial thickness for different patients and for different oral sites, all data in Fig. 6 were normalized by the corresponding mean ensemble value obtained from an adjacent or contralateral normal reference that was similarly acquired for each patient.

The data in Fig. 6 reveal a significant increase in depth-dependent epithelial backscattering for almost all imaged dysplastic tissue sites with a biopsy-confirmed diagnosis. Lichenoid mucositis (LM, first column in Fig. 6) without obvious dysplasia is a benign lesion of the oral cavity that showed similar backscattering ratio as the corresponding normal reference. We observe the relative increase in the backscattering ratio to be different for different biopsy sites with the same histopathological diagnosis. However, within the same anatomical site (see data for ventral tongue in Fig. 6), we find a trend toward increasing backscattering ratio when the degree of pathology increases.

The provided minimum and maximum values (whiskers in Fig. 6) indicate a large variability within each frame ensemble which is caused by small intended and unintended motion of the imaging wand about the heterogeneous biopsy area. We believe that variations in epithelial thickness have a very strong influence on our computed backscattering ratios.<sup>27</sup> This may explain the opposite data trends for the two mild dysplasia biopsies (second and third column in Fig. 6) and the differences found between anatomical sites.<sup>26</sup>

One may argue the influence of fluorescence staining on measured reflectance values. To study this effect, we performed multiple measurements in a volunteer's mouth at the buccal mucosa both before and after topical staining with AH. In

this test, we found a 3% to 6% influence on ratio backscattering. Absorption effects in the stained surface cell layers impact incident blue light twice, once upon entering and once upon exiting the medium. This diminishing influence tends to be proportional to the spatial frequency-dependent reflectance value and is thus reduced by computing reflectance ratios.

In order to ensure a wetted contact between the fiber imaging probe and the tissue surface, a minimum contact pressure is required. According to our observations, further increase in contact pressure (i.e., tissue squeezing) has little influence on quantified backscattering ratios (less or equal to 5%).

In the next analysis, we perform quantitative data analysis of both imaging modes by comparing the nuclear-to-cytoplasmic ratio (NCR, i.e., the perceived nuclear area density) in FE to averaged epithelial backscattering in DOM. This will allow for identifying the main malignancy-associated changes in epithelial backscattering in dependence of the degree of epithelial dysplasia.

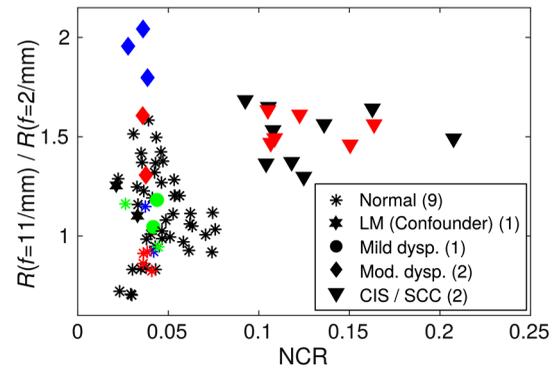
To reduce the confounding factor of epithelial thickness to backscattering signals, we restrict our evaluation to image frames from lesions and normal references of the lateral tongue, ventral tongue, and floor of the mouth. Furthermore, we limit our analysis to data for which a known pathology-based diagnosis is available and in which the NCR is quantifiable in FE at the same time. In spite of expected differences in epithelial thickness for the lateral tongue and floor of the mouth,<sup>26</sup> we do not further subdivide our data evaluation because of the limited number of cases in this study.

Quantitative evaluation of FE images is manually performed by counting the number density and by measuring the mean size of nuclei with the image analysis tool ImageJ. Multiplication of these two values yields an approximation of the NCR as given in Fig. 7. The second axis of data presented in this graph quantifies the backscattering ratio averaged over the central imaging field of view for each frame.

While the number of lesions considered for Fig. 7 is limited, we can still observe an increase in the backscattering ratio as the epithelium transforms from a normal state toward advanced dysplasia or cancer. For this change in backscattering ratio to occur, the moderately or severely dysplastic epithelium has to exhibit either elevated  $R(f = 11 \text{ mm}^{-1})$  values and/or lower  $R(f = 2 \text{ mm}^{-1})$  values. This corresponds to more pronounced backscattering from the superficial epithelial region and/or less scattering from the lower epithelial region.

It is revealing to look at the data from the three patients that contribute data to multiple diagnostic groups which we separate by red, green, and blue color in Fig. 7, respectively. From these data points, we learn that the relative change in backscattering ratio is much more significant when evaluated for a single patient over a range of pathologies. The red data points indicate a fairly constant backscattering ratio as the epithelium progresses from moderate to severe dysplasia which is contrary to the seeming decline when looking at all data points. Owing to the same imaging field of view for both modalities, the separation between severe dysplasia and all other data points is achieved with very high certainty, as nuclear density serves as an established imaging feature to confirm severely dysplastic lesions.<sup>11</sup> This removes some of the localization uncertainty often encountered for other imaging devices when coregistering images to histopathology.

As a final step in our data evaluation and to further confirm the correlation between depth-dependent backscattering and

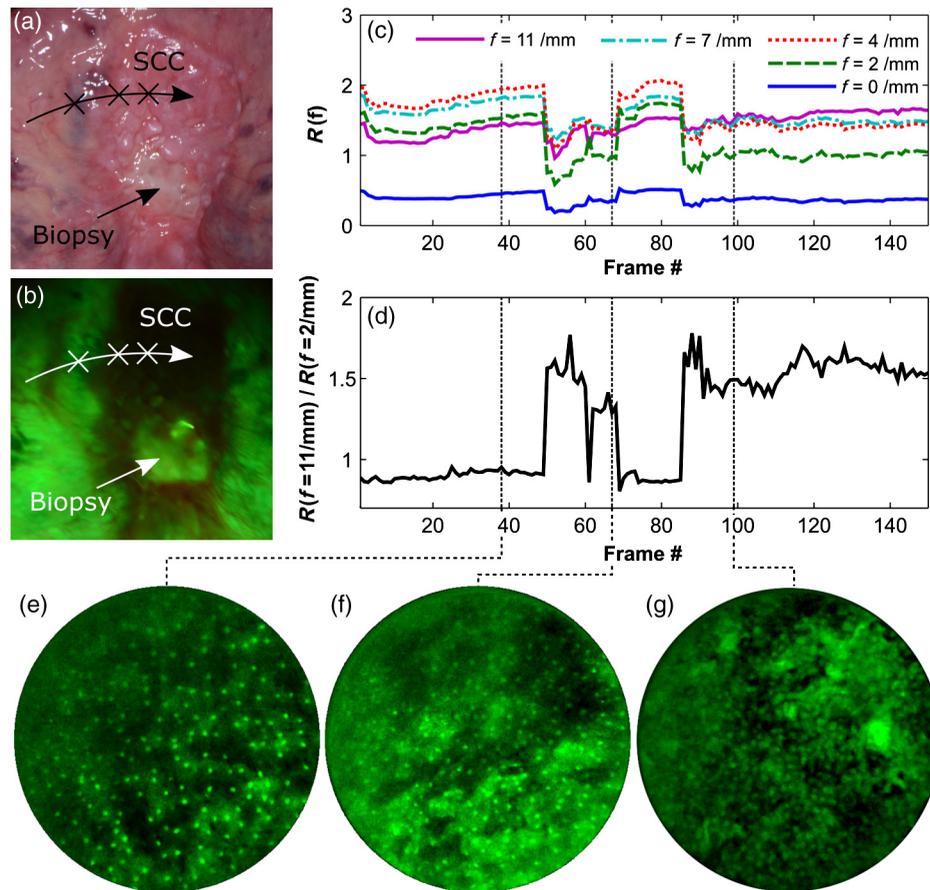


**Fig. 7** Simultaneous evaluation of the NCR using FE and of the backscattering ratio using DOM. Every data point corresponds to an independent frame captured at a different tissue site of either the ventral or lateral tongue or the floor of mouth. The presented black data points correspond to multiple measurements from various sites in eight patients. We distinguish the data from three selected patients in green, blue, or red, as each patient contributes data to more than one diagnostic group. The legend numbers state how many patients contribute to each data group. Data points for normal epithelium comprise imaging on both patients and on a healthy volunteer (30 years).

dysplasia, we provide a more in depth analysis of one particular lesion. In images a and b of Fig. 8, we present clinical white light and fluorescence visualization (FV) photographs of a SCC on the ventral tongue of a patient, respectively.<sup>29</sup> For this lesion, the clinician used our imaging wand to scan from a normal to an abnormal area [see arrows in Figs. 8(a) and 8(b)]. During this scan of about 40 s, our device roughly recorded 150 frames without motion artifacts for which we quantify averaged reflectance intensities and the reflectance ratio in Figs. 8(c) and 8(d), respectively.

Absolute reflectance values in Fig. 8(c) quantify tissue backscattering throughout this scan relative to our intralipid flat-field reference. As a first observation, reflectance values at zero spatial frequency [ $R(f = 0 \text{ mm}^{-1})$ ] are found to mostly lie below that of our reference ( $R < 1$ ), whereas reflectance at all other spatial frequencies mostly exceeds that of the intralipid reference ( $R > 1$ ). This indicates that backscattering from the epithelium is stronger than that of the connective tissue below. Analysis of the top four curves in Fig. 8(c) reveals characteristic changes in the relative behavior of epithelial reflectance. While  $R(f = 11 \text{ mm}^{-1})$  is below the other three curves for data from the normal tissue area (frames 0 to 48), it demonstrates the strongest signal among the four curves in the SCC area (high frame numbers near end of scan). This flip in depth-dependent epithelial backscattering becomes particularly apparent in the corresponding backscattering ratio in Fig. 8(d). Three FE images [Figs. 8(e), 8(f), and 8(g)] show the nuclear pattern for the frames, which are marked in the data graphs (dotted lines) and in the clinical photographs (crosses).

We attribute the repeated inversion of reflectance values in the transition zone (frames 49 to 84) to the heterogeneity of the lesion boundary. Interestingly, the second FE image [Fig. 8(f) corresponding to frame 66] shows no nuclear crowding, whereas the reflectance ratio [Fig. 8(d)] already indicates significant changes in epithelial backscattering. We explain this by the sensitivity of DOM to non-full-thickness tissue alterations consistent with mild and moderate dysplasia, which also agrees with the loss of fluorescence in the FV image [Fig. 8(b)].



**Fig. 8** In-depth analysis of a DME scan from a normal area into a SCC of the ventral tongue. Clinical white light and fluorescence validation photographs (a and b) illustrate the scan path. Reflectance values for the corresponding frames (c) reveal a shift in depth-dependent scattering as highlighted in the backscattering ratio (d). (e)–(g) show FE images corresponding to both the vertical lines in (c) and (d) and to the three crosses in (a) and (b).

In order to reduce motion artifacts, the tissue scan of Fig. 8 was not performed at constant speed but rather in a stop-and-go pattern. This explains the abrupt changes in reflectance as presented in Figs. 8(c) and 8(d).

#### 4 Discussion

It should be emphasized, that we do not observe the drop in absolute epithelial backscattering [Fig. 8(c)] to be a specific feature of a dysplastic epithelium but rather the relative depth-dependent shift in backscattering as quantified by the reflectance ratio [Fig. 8(d)]. Even though we previously found in Sec. 3.1 that keratin seems to play a major role in tissue backscattering [see Fig. 5(c)] we did not observe the depth-dependent reflectance ratios to be significantly influenced by hyperkeratosis. As an example, the keratin-related scattering contrast presented in Fig. 5(c) is mostly eliminated by forming reflectance ratios for the different spatial frequencies. Similarly, out of the two patients with leukoplakia (i.e., a clinical diagnosis for keratin-induced tissue whitening) only one showed an elevated backscattering ratio.

In strongly vascularized imaging areas, we notice that the associated decrease in reflectance is relatively more pronounced for low spatial frequencies than for high spatial frequencies. This effect alters computed backscattering ratios to some extent. Our evaluation disregards this effect by averaging over

these areas thereby improving signal-to-noise ratios. A more advanced evaluation would have to take these depth-dependent absorption effects into account or make use of red or infrared light sources to avoid tissue absorption altogether. Our choice of the blue light source was motivated by the obtained vascular imaging contrast. However, until now we were not able to correlate vascular patterns in DOM to specific tissue diagnoses.

In the approach presented herein, a single blue light source is used to run both imaging modalities. We notice only a small interference between the two imaging modes. Nevertheless, higher precision in the obtained reflectance values in the absence of vascular contrast (blood absorption of the illumination light) could be obtained by synchronous tissue illumination with a second wavelength for absorption-independent measurement of tissue backscattering.

The extent of this study is too small to conclude much about sensitivity and specificity values for DME in the oral cavity. Nevertheless, we were able to demonstrate complementary contrast of the two imaging modalities and anticipate a correspondingly high diagnostic yield.

#### 5 Conclusions

The design of our DME system is tailored to provide noninvasive and real-time sensitivity to identify oral lesions at risk. Through the parallel use of two different imaging modalities

that provide contrast for different depth regimes, we achieve sensitivity to both moderate and severe dysplasia. FE can image the nuclear pattern of the uppermost cell layer and is thus sensitive to severe dysplasia where structural changes extend to the epithelial surface. In the majority of oral lesions, excessive fluorescence from cytoplasmic keratin strongly distorts FE contrast.<sup>30</sup> DOM can improve the limited diagnostic yield of FE and may also be sensitive to earlier forms of precancer by quantifying depth-dependent epithelial backscattering changes.

Our studies indicate a strong shift in depth-dependent epithelial backscattering related to the development of precancer.<sup>5</sup> We find this shift, rather than absolute changes to be strongly correlated with epithelial dysplasia. Our findings from this pilot study confirm both our previous DOM analysis of the uterine cervix and independent results from two OCT studies in the oral cavity.<sup>28,24</sup> Similarly, the demonstrated significance of high spatial frequency reflectance confirms previous findings using spatially resolved reflectance where very short source-detector separations yielded highest diagnostic sensitivities.<sup>31,32</sup> It is yet to be determined if shifts in depth-dependent epithelial backscattering prove to be a characteristic feature of precancerous development in the epithelium independent of organ site.

The total hardware costs of our DME system in volumes of one amount to about 8000 USD. Given this relatively low cost and in anticipation of our results being confirmed in a more comprehensive patient study, we believe DME may serve as a valuable tool for both biopsy guidance and *in vivo* diagnosis and possibly also for tumor margin assessment.

In a practical use case, clinicians would point the imaging wand to abnormal tissue sites that were previously stained with AH. Either FE or DOM or both may then sense high-grade dysplasia and trigger biopsy collection to confirm diagnosis. The added benefit of DOM mostly lies in the real-time provision of computed backscattering ratio values as a measure for tissue abnormality.

As future improvements to our system, we plan to increase the light source intensity and thereby enhance the obtained signal-to-noise ratio. Until now, we refrained from using spatial frequencies lower than  $f = 2 \text{ mm}^{-1}$  because of the inability to accurately measure low spatial frequency reflectance with a fiber bundle of diameter  $790 \mu\text{m}$ . Proper selection of spatial frequencies, optimized signal quality, and robust data evaluation strategies are key to maximize diagnostic sensitivity. This could allow one to overcome the major challenge for DOM, which is the natural variability in epithelial thickness for different imaging sites.

### Disclosures

No conflicts of interest, financial or otherwise, are declared by the authors.

### Acknowledgments

We acknowledge project funding by grants P01 CA82710 of the National Institutes of Health (NIH) and CPG-134740 of the Canadian Institutes of Health Research (CIHR). We strongly appreciate the participation of all 21 patients in this study. Personal thanks go to Tony Han, Anthony Lee, Rashika Raizada, and the dental assistants of the BC Cancer Agency Dental Clinic.

### References

1. B. W. Neville and T. A. Day, "Oral cancer and precancerous lesions," *CA Cancer J. Clin.* **52**(4), 195–215 (2002).
2. S. Shirani et al., "Epithelial dysplasia in oral cavity," *Iran J. Med. Sci.* **39**(5), 406–417 (2014).
3. A. Moscicki et al., "Rate of and risks for regression of CIN-2 in adolescents and young women," *Obstet. Gynecol.* **116**(6), 1373–1380 (2010).
4. M. Contaldo et al., "Oral mucosa optical biopsy by a novel handheld fluorescent confocal microscope specifically developed: technologic improvements and future prospects," *Oral Surg. Oral Med. Oral Pathol. Oral Radiol.* **116**(6), 752–758 (2013).
5. N. Bodenschatz et al., "Diffuse optical microscopy for quantification of epithelial backscattering in the cervix," *J. Biomed. Opt.* **21**(6), 066001 (2016).
6. A. L. Polglase et al., "A fluorescence confocal endomicroscope for *in vivo* microscopy of the upper- and the lower-GI tract," *Gastrointest. Endosc.* **62**(5), 686–695 (2005).
7. M. K. Quinn et al., "High-resolution microendoscopy for the detection of cervical neoplasia in low-resource settings," *PLoS One* **7**(9), e44924 (2012).
8. P. M. Vila et al., "Discrimination of benign and neoplastic mucosa with a high-resolution microendoscope (HRME) in head and neck cancer," *Ann. Surg. Oncol.* **19**(11), 3534–3539 (2012).
9. P. S. Thong et al., "Review of confocal fluorescence endomicroscopy for cancer detection," *IEEE J. Sel. Top. Quantum Electron.* **18**(4), 1355–1366 (2012).
10. P. M. Lane et al., "Confocal fluorescence microendoscopy of bronchial epithelium," *J. Biomed. Opt.* **14**(2), 024008 (2009).
11. C. Schlosser et al., "Fluorescence confocal endomicroscopy of the cervix: pilot study on the potential and limitations for clinical implementation," *J. Biomed. Opt.* **21**(12), 126011 (2016).
12. D. Shin et al., "Quantitative analysis of high-resolution microendoscopic images for diagnosis of esophageal squamous cell carcinoma," *Clin. Gastroenterol. Hepatol.* **13**, 272–279.e2 (2015).
13. T. J. Muldoon et al., "Noninvasive imaging of oral neoplasia with a high-resolution fiber-optic microendoscope," *Head Neck* **34**(3), 305–312 (2012).
14. J. Tan et al., "Detection of cervical intraepithelial neoplasia *in vivo* using confocal endomicroscopy," *Gynaecol. Oncol.* **166**, 1663–1670 (2009).
15. P. M. Lane, "Terminal reflections in fiber-optic image guides," *Appl. Opt.* **48**(30), 5802–5810 (2009).
16. B. Obstoy et al., "Safety and performance analysis of acriflavine and methylene blue for *in vivo* imaging of precancerous lesions using fibered confocal fluorescence microscopy (FCFM): an experimental study," *BMC Pulm. Med.* **15**(30), 1–10 (2015).
17. N. Bodenschatz et al., "Quantifying phase function influence in subdiffusively backscattered light," *J. Biomed. Opt.* **21**(3), 035002 (2016).
18. P. A. Keahey et al., "Optimizing modulation frequency for structured illumination in a fiber-optic microendoscope to image nuclear morphology in columnar epithelium," *Biomed. Opt. Express* **6**(3), 870–880 (2015).
19. P. Keahey et al., "Differential structured illumination microendoscopy for *in vivo* imaging of molecular contrast agents," *Proc. Natl. Acad. Sci. U.S.A.* **113**, 10769–10773 (2016).
20. N. Bozinovic et al., "Fluorescence endomicroscopy with structured illumination," *Opt. Express* **16**(11), 8016–8025 (2008).
21. J. G. Batsakis, "Clinical pathology of oral cancer," in *Oral Cancer*, J. P. Shah, J. G. Batsakis, and N. W. Johnson, p. 87, Dunitz Martin Ltd., London (2003).
22. N. Bodenschatz et al., "Surface layering properties of intralipid phantoms," *Phys. Med. Biol.* **60**, 1171–1183 (2014).
23. S. El Hallani et al., "Ex vivo confocal imaging with contrast agents for the detection of oral potentially malignant lesions," *Oral Oncol.* **49**, 582–590 (2013).
24. M.-T. Tsai et al., "Differentiating oral lesions in different carcinogenesis stages with optical coherence tomography," *J. Biomed. Opt.* **14**(4), 044028 (2009).
25. A.-M. Pena et al., "Spectroscopic analysis of keratin endogenous signal for skin multiphoton microscopy," *Opt. Express* **13**(16), 6268–6274 (2005).
26. S. Prestin et al., "Measurement of epithelial thickness within the oral cavity using optical coherence tomography," *Head Neck* **34**, 1777–1781 (2012).

27. C. Arens et al., "Histologic assessment of epithelial thickness in early laryngeal cancer or precursor lesions and its impact on endoscopic imaging," *Eur. Arch. Otorhinolaryngol.* **264**, 645–649 (2007).
28. O. K. Adegun et al., "Quantitative analysis of optical coherence tomography and histopathology images of normal and dysplastic oral mucosal tissues," *Lasers Med. Sci.* **27**, 795–804 (2012).
29. P. M. Lane et al., "Simple device for the direct visualization of oral-cavity tissue fluorescence," *J. Biomed. Opt.* **11**(2), 024006 (2006).
30. T. Collier et al., "Real-time reflectance confocal microscopy: comparison of two-dimensional images and three-dimensional image stacks for detection of cervical precancer," *J. Biomed. Opt.* **12**(2), 024021 (2007).
31. N. M. Marin et al., "Diffuse reflectance patterns in cervical spectroscopy," *Gynecol. Oncol.* **99**, S116–S120 (2005).
32. Y. N. Mirabal et al., "Reflectance spectroscopy for in vivo detection of cervical precancer," *J. Biomed. Opt.* **7**(4), 587–594 (2002).

Biographies for the authors are not available.