

## Development of an integrated multimodal optical imaging system with real-time image analysis for the evaluation of oral premalignant lesions

Eric C. Yang  
Imran S. Vohra  
Hawraa Badaoui  
Richard A. Schwarz  
Katelin D. Cherry  
Timothy Quang  
Justin Jacob  
Alex Lang  
Nancy Bass  
Jessica Rodriguez  
Michelle D. Williams  
Nadarajah Vigneswaran  
Ann M. Gillenwater  
Rebecca R. Richards-Kortum

Eric C. Yang, Imran S. Vohra, Hawraa Badaoui, Richard A. Schwarz, Katelin D. Cherry, Timothy Quang, Justin Jacob, Alex Lang, Nancy Bass, Jessica Rodriguez, Michelle D. Williams, Nadarajah Vigneswaran, Ann M. Gillenwater, Rebecca R. Richards-Kortum, "Development of an integrated multimodal optical imaging system with real-time image analysis for the evaluation of oral premalignant lesions," *J. Biomed. Opt.* **24**(2), 025003 (2019), doi: 10.1117/1.JBO.24.2.025003.

# Development of an integrated multimodal optical imaging system with real-time image analysis for the evaluation of oral premalignant lesions

Eric C. Yang,<sup>a,b</sup> Imran S. Vohra,<sup>a</sup> Hawraa Badaoui,<sup>c</sup> Richard A. Schwarz,<sup>a</sup> Katelin D. Cherry,<sup>a</sup> Timothy Quang,<sup>a</sup> Justin Jacob,<sup>c</sup> Alex Lang,<sup>d</sup> Nancy Bass,<sup>d</sup> Jessica Rodriguez,<sup>c</sup> Michelle D. Williams,<sup>e</sup> Nadarajah Vigneswaran,<sup>d</sup> Ann M. Gillenwater,<sup>c</sup> and Rebecca R. Richards-Kortum<sup>a,\*</sup>

<sup>a</sup>Rice University, Department of Bioengineering, Houston, Texas, United States

<sup>b</sup>Baylor College of Medicine, MD/PhD Medical Scientist Training Program, Houston, Texas, United States

<sup>c</sup>University of Texas M.D. Anderson Cancer Center, Department of Head and Neck Surgery, Houston, Texas, United States

<sup>d</sup>University of Texas Health Science Center, School of Dentistry, Department of Diagnostic and Biomedical Sciences, Houston, Texas, United States

<sup>e</sup>University of Texas M.D. Anderson Cancer Center, Department of Pathology, Houston, Texas, United States

**Abstract.** Oral premalignant lesions (OPLs), such as leukoplakia, are at risk of malignant transformation to oral cancer. Clinicians can elect to biopsy OPLs and assess them for dysplasia, a marker of increased risk. However, it is challenging to decide which OPLs need a biopsy and to select a biopsy site. We developed a multimodal optical imaging system (MMIS) that fully integrates the acquisition, display, and analysis of macroscopic white-light (WL), autofluorescence (AF), and high-resolution microendoscopy (HRME) images to noninvasively evaluate OPLs. WL and AF images identify suspicious regions with high sensitivity, which are explored at higher resolution with the HRME to improve specificity. Key features include a heat map that delineates suspicious regions according to AF images, and real-time image analysis algorithms that predict pathologic diagnosis at imaged sites. Representative examples from ongoing studies of the MMIS demonstrate its ability to identify high-grade dysplasia in OPLs that are not clinically suspicious, and to avoid unnecessary biopsies of benign OPLs that are clinically suspicious. The MMIS successfully integrates optical imaging approaches (WL, AF, and HRME) at multiple scales for the noninvasive evaluation of OPLs. © The Authors. Published by SPIE under a Creative Commons Attribution 4.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: 10.1117/1.JBO.24.2.025003]

Keywords: oral cancer; oral lesion; *in vivo* imaging; autofluorescence; fiber bundle.

Paper 180600R received Oct. 31, 2018; accepted for publication Jan. 30, 2019; published online Feb. 21, 2019.

## 1 Introduction

With over 300,000 new cases per year and a mortality rate of ~50%, oral cancer is a major global health issue.<sup>1</sup> The stage at diagnosis is the most important predictor of survival, and unfortunately, most patients are diagnosed at a late stage. Therefore, earlier diagnosis of oral cancer and its precursor lesions is critical to improved outcomes. Most oral cancers originate as oral premalignant lesions (OPLs), defined as oral mucosal lesions with elevated risk of malignant transformation including leukoplakia and erythroplakia. OPLs affect millions worldwide and are challenging for clinicians to identify and manage, contributing to late-stage diagnosis. Some guidelines recommend that all suspicious lesions receive an initial biopsy<sup>2</sup> to assess the OPL for dysplasia, the most well-established risk marker of malignant transformation. Dysplasia can be graded as mild, moderate, or severe, and the risk of malignant transformation increases with grade.<sup>3</sup>

This management protocol is not optimal. Clinicians must distinguish OPLs from benign confounder lesions with a similar appearance, a challenging task for general practitioners, who are most likely to initially evaluate patients with OPLs. Then, the clinician must select a biopsy site within the lesion and perform the biopsy with proper technique. Only a small percentage of

OPLs contain dysplasia, leading to unnecessary biopsies. Biopsies are highly invasive, resource intensive, and require days to process and interpret. For these reasons, many OPLs are not biopsied, particularly in low-resource settings with the highest prevalence of OPLs, such as South Asia. Additionally, the presence and severity of dysplasia often vary within an OPL. Clinicians try to biopsy the tissue with the highest grade of pathology, but are unsuccessful up to 30% of the time.<sup>4</sup> Finally, there are no clear guidelines on when to obtain additional biopsies in patients with lesions under surveillance for potential progression or with a history of oral cancer being monitored for recurrence.

A noninvasive, point of care diagnostic adjunct capable of distinguishing dysplastic or cancerous lesions from benign lesions and guiding clinicians to the optimal biopsy site could decrease unnecessary biopsies and reduce underdiagnosis. Existing adjuncts, including brush biopsy, toluidine blue, acetowhitening, and autofluorescence imaging (AF) do not have sufficient accuracy or, in the case of brush biopsy, do not provide results at the point of care.<sup>2,5</sup>

AF is based on the premise that dysplastic progression leads to alterations in the concentration of native tissue fluorophores, including NADH, FAD, collagen, elastin, and protoporphyrin IX, and altered light scattering due to changes in epithelial thickness and nuclear morphology. Empirically, dysplastic tissue exhibits a large loss in green autofluorescence (driven primarily by reduction of collagen fluorescence) and a small increase in red autofluorescence (driven by endogenous porphyrins).<sup>6-8</sup>

\*Address all correspondence to Rebecca R. Richards-Kortum, E-mail: rkortum@rice.edu

Perceptually, the fluorescence of nonsuspicious tissue appears bright and suspicious tissue appears dark (known as “loss of fluorescence”). AF quickly assesses large fields of mucosa and detects dysplasia and cancer with high sensitivity, but its clinical utility is hampered by low specificity for benign inflammatory lesions, which comprise the majority of oral mucosal lesions presenting to general dental practitioners.<sup>9–11</sup>

We developed a device called the high-resolution microendoscope (HRME) that has higher specificity than AF but only assesses 0.5 mm<sup>2</sup> of mucosa at a time.<sup>12,13</sup> The HRME is a fluorescence microscope coupled to a 0.79-mm-diameter optical fiber bundle. The distal tip of the fiber is placed in contact with the mucosa after topical application of the fluorescent dye proflavine. Proflavine stains nuclei, enabling the HRME to visualize nuclear morphology in the superficial epithelium, which is altered in high-grade dysplasia and cancer but not in benign inflammatory lesions. Previously, we imaged patients with oral lesions in both high- and low-risk populations using separate AF and HRME imaging systems and developed automated algorithms to calculate image features [normalized red to green (RG) ratio from AF and number of abnormal nuclei/mm<sup>2</sup> from HRME] that retrospectively distinguished high-grade dysplasia and cancerous sites from benign sites with high accuracy.<sup>14–16</sup>

Based on these results, we propose a two-step imaging procedure to evaluate OPLs that combine the individual strengths of AF and HRME. First, AF is used to identify high-risk regions within the lesions with high sensitivity. Then, those regions are explored with the HRME to reduce false positives. This procedure could identify dysplastic or cancerous lesions and guide the clinician to the optimal biopsy site. To implement this procedure, methods must be developed to quickly identify

high-risk regions by AF, alert clinicians to those regions, find those regions in the patient’s mouth, and precisely locate HRME image sites on AF images so that diagnostic information from both modalities can be combined.

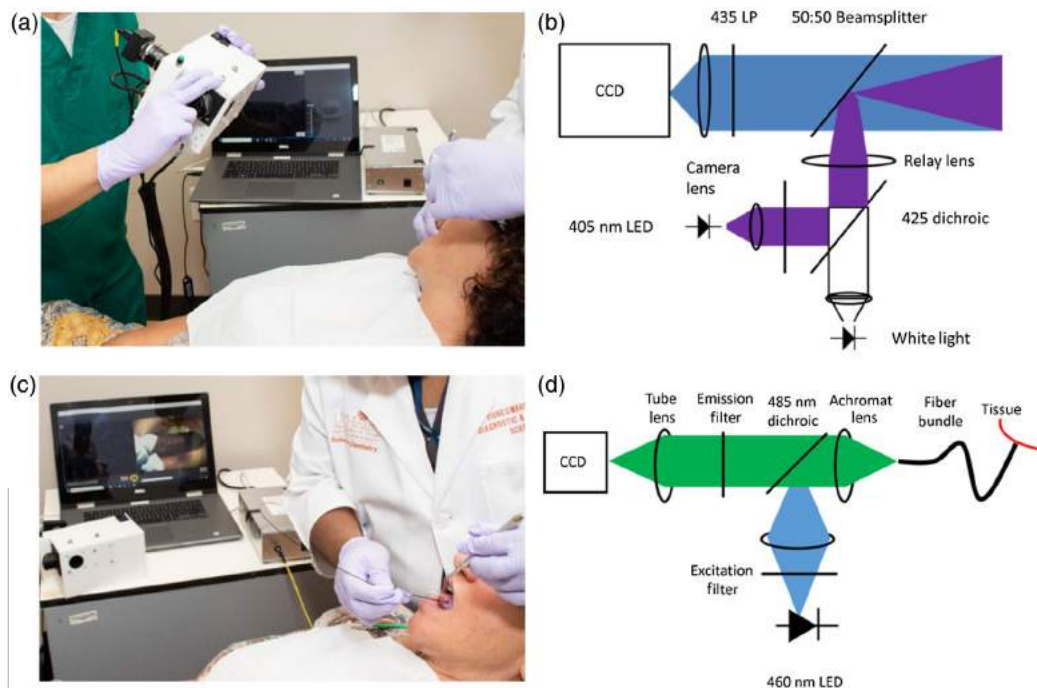
In this paper, we develop a multimodal optical imaging system (MMIS) that fully integrates AF and HRME together with real-time image analysis. Key features include a heat map that delineates suspicious regions according to AF, an image registration algorithm that aligns AF and white-light images, and an efficient procedure to mark the location of HRME images. We then present representative results from ongoing clinical studies of the MMIS that demonstrate its potential to improve care for patients with oral lesions.

## 2 Materials and Methods

### 2.1 Instrumentation

The MMIS acquires widefield white-light reflectance (WL), widefield autofluorescence (AF), and HRME images. The instrument is connected by universal serial bus (USB) to a consumer touchscreen laptop that runs the multimodal imaging software. Arduino Nano microcontrollers are used for electronic control, and power is supplied through rechargeable lithium batteries and USB. The MMIS can be built with parts costing ~\$3000.

Widefield WL and AF images are acquired as shown in Fig. 1(a) with the optical setup shown in Fig. 1(b). AF images are acquired using 405-nm LED excitation light that passes through a 400-/40-nm bandpass filter, is reflected across a 425-nm dichroic mirror, and is focused on a beamsplitter before tissue excitation. Fluorescence emission passes through a 435-nm longpass filter and is focused onto a color CCD (Point Gray



**Fig. 1** Multimodal imaging system. (a) Acquisition of widefield WL and AF images of a patient’s OPL. Widefield acquisition occurs with the room lights off; for visualization purposes, the lights were left on. The other MMIS components (touchscreen laptop and HRME instrumentation) are visible in the background. (b) Schematic diagram of widefield WL and AF optical instrumentation. (c) Acquisition of HRME images of a patient’s OPL. The tip of the fiber probe is placed in gentle contact with the OPL after topical application of proflavine dye. (d) Schematic diagram of HRME optical instrumentation.

CMLN-13S2C-CS), which saves eight-bit RGB images. WL images are acquired with a white illumination LED. The resulting images have a 4.5-cm-diameter field of view and 100- $\mu$ m lateral resolution. Three manual controls are available on the instrument: a switch to toggle illumination on or off, a switch to alternate between WL and AF illumination, and a button to initiate an image-acquisition sequence.

HRME images are acquired as shown in Fig. 1(c) using the optical setup shown in Fig. 1(d). Briefly, excitation light from a 460-nm LED passes through a 475-nm shortpass filter, is reflected across a 485-nm dichroic mirror, and is focused onto a 790- $\mu$ m-diameter multimodal optical fiber bundle with 4.4- $\mu$ m core-to-core spacing. Fluorescence emission travels through the fiber bundle, dichroic, 500-nm longpass filter, and is focused onto a monochrome CCD (Point Gray CMLN-13S2M-CS). To acquire HRME images, the user places the tip of the fiber in gentle contact with the oral mucosa after topical application of sterile proflavine (0.01% w/v in PBS), a fluorescent dye that stains cell nuclei, with a cotton-tipped applicator. A foot pedal can freeze and unfreeze the image feed, and a switch on the instrument toggles the illumination on or off.

## 2.2 Multimodal Imaging Procedure and Software

The MMIS software provides a user interface that fully integrates the acquisition, display, and real-time image analysis of WL, AF, and HRME images. Specifically, the MMIS software directs the clinician through the following steps:

1. Acquire widefield WL and AF images of the lesion(s).
2. Outline suspicious regions on touchscreen based on clinical impression and the processed AF image.
3. Explore suspicious regions with HRME, saving images at desired locations.
4. Identify locations of saved HRME images on touchscreen.
5. Review full multimodal imaging results.

These steps are detailed below and demonstrated in Fig. 2 (Video 1), a video that shows selected screens of the touchscreen laptop during use of the MMIS. Imaging is typically completed in 10 min.



**Fig. 2** Selected screens during use of the MMIS (Video 1, MP4, 26.5 MB [URL: <https://doi.org/10.1117/1.JBO.24.2.025003.1>]).

### 2.2.1 Acquire widefield white-light and autofluorescence images of the lesion(s)

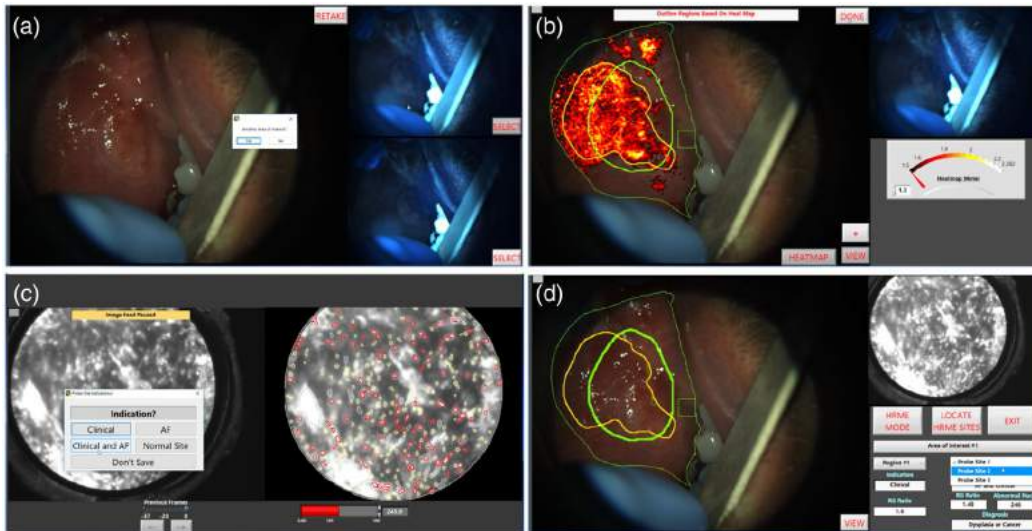
The first step is to acquire widefield WL and AF images of the lesion(s). Lesions are imaged using the handheld instrumentation with room lights off, with another participant manually exposing the lesion [Fig. 1(a)]. When a satisfactory image frame is achieved, the user presses the acquisition button to save one WL image and two AF images with different preset exposure and gain settings. A single-acquisition sequence lasts 0.94 s. A live focus bar helps the user acquire in-focus images (see “Focus metric” section of the Sec. 2.3); in the case of motion blur due to movement from the user or the patient, the acquisition can be repeated. Multiple acquisitions can be performed if the patient has multiple lesions. The software displays the three images, and one of the two AF images is chosen by the user based on good FOV (lesion centered and adjacent normal mucosa visible), lack of motion blur, signal intensity within dynamic range, and good focus [Fig. 3(a)]. Background images without LED illumination are also acquired with the same exposure and gain settings as the AF images as part of the acquisition sequence and are automatically subtracted to eliminate the contribution of remaining ambient light. The clinician is then prompted to outline the mucosa on the WL image using the touchscreen to exclude nonmucosal objects such as teeth or retractors.

### 2.2.2 Outline regions based on clinical suspicion and autofluorescence heatmap on touchscreen

In step two, the clinician outlines the most clinically suspicious regions within the lesion(s) on the WL image using the touchscreen laptop. Meanwhile, the WL image and selected AF image are processed. First, an image registration algorithm aligns the WL and AF images. The results are used to convert the outline of the mucosa (step 1) from WL to AF coordinates, which is used to calculate the normalized RG ratio at each pixel in the mucosa. The normalized RG ratio values are used to generate a heat map, which is converted from AF to WL coordinates and overlaid on the WL image displayed by the software. See Sec. 2.3 (“Automated normalized RG ratio calculation, Heat map generation, and Widefield image registration algorithm” subsections) for technical details of these image processing algorithms.

The clinician then outlines suspicious regions based on the heat map [Fig. 3(b)]. To facilitate this, the software has three controls to adjust the display. The first control (“view” button) alternates the heat map overlay between the WL and AF images. The second control is a meter that adjusts the minimum normalized RG ratio of pixels highlighted by the heat map, allowing the clinician to adjust the AF threshold in real time. The third control (“heatmap” button) toggles the heat map on or off.

The software then moves to the summary screen, an interactive display summarizing the multimodal imaging results. At this point, the summary screen shows the acquired images, outlined regions, and image analysis results. A widefield image (WL or AF) is displayed in the main image display on the left along with the outlined regions. A “view” button alternates between WL and AF in the image display. The bottom right of the summary screen contains dropdown menus and image analysis results. If multiple widefield acquisitions were performed in step one, a dropdown menu allows the clinician to select the particular acquisition to view results. A second dropdown



**Fig. 3** MMIS software during multimodal imaging of an OPL patient. For improved visualization, the brightness of panels (a), (b), and (d) has been increased. (a) WL image (left) and two AF images (right) acquired during a single acquisition sequence. Dialog box allows the user to image additional lesions. (b) AF-based heat map overlaid on the WL image (left); suspicious regions based on the heat map (yellow outline) and clinical suspicion (green outline) are visible. The AF image is displayed for reference (top right). The heatmap meter (bottom right) can be adjusted so that the heat map highlights fewer pixels. (c) The HRME image (left) after image analysis (right). The meter (bottom right) displays the number of abnormal nuclei/mm<sup>2</sup>. (d) Summary screen following imaging procedure. The WL image is displayed (left) with suspicious regions and HRME site locations overlaid. Two dropdown menus are available to select a suspicion region or a probe site and view its normalized RG ratio, HRME risk score, and predicted diagnosis (bottom right). The HRME image corresponding to the selected probe site is displayed in the top right.

menu allows the clinician to select an outlined region and view its normalized RG ratio, indication (WL or AF), and location.

### 2.2.3 Explore outlined regions with high-resolution microendoscopy

Step three is to explore the suspicious regions outlined in step two with the HRME probe. The user presses the “HRME mode” button on the summary screen and applies proflavine to the mucosa. The left image display shows the live HRME image feed, and the right image shows a widefield image—the “view” button alternates between WL and AF—with the outlined regions overlaid for easy reference. The clinician explores each outlined region with the HRME probe and presses the foot pedal to pause the feed at sites of interest. If the paused frame has motion artifact, the clinician can use arrow buttons to scroll through the 37 most recent frames, then press the button with the save symbol to select the image. The image analysis algorithm calculates the number of abnormal nuclei/mm<sup>2</sup> in the selected image (see Sec. 2.3.2 of Sec. 2.3) and displays the result [Fig. 3(c)]. If the clinician decides to save an image from a site, they then make an ink dot with a marking pen on the patient’s mucosa at the HRME site and can designate the indication for the image. After HRME images have been acquired to the clinician’s satisfaction, the “back” button is pressed to return to the summary screen.

### 2.2.4 Locate high-resolution microendoscopy sites on touchscreen

The goal of step four is to locate the HRME sites on the original AF images using the ink dots. From the summary screen, the

clinician presses the “locate HRME sites” button and acquires another WL image, of similar orientation to the initial WL image, which includes the ink dots. The initial WL image is shown in the right image display for reference, and the focus bar helps ensure an in-focus image. After acquisition, the “ink dot” WL image is displayed on the right, and the initial WL image is displayed on the left. The clinician marks the HRME sites on the initial WL image, using the ink dot image as a reference, with the touchscreen. The HRME sites are converted from WL to AF coordinates to calculate their normalized RG ratio, based on a 17-pixel diameter circle. The normalized RG ratio at each HRME site is combined with its number of abnormal nuclei/mm<sup>2</sup> for classification as benign or moderate dysplasia-cancer, using  $(47.1 \times \text{normalized RG ratio}) + (\text{number of abnormal nuclei/mm}^2) = 273.7$  as a linear decision boundary (see Sec. 2.3.3 of Sec. 2.3).

### 2.2.5 Review full multimodal imaging results

The full multimodal imaging results are then interactively displayed in the summary screen [Fig. 3(d)]. In addition to the outlined regions, the widefield image in the main image display is overlaid with circles representing sites where HRME images were obtained. Individual HRME sites can be selected from a dropdown menu to view the image (top right), its location, indication, normalized RG ratio, number of abnormal nuclei/mm<sup>2</sup>, and predicted diagnosis.

The “exit” button closes the software and automatically saves a text file summarizing the MMIS session. The text file lists the imaged lesions, the selected AF image, the indication and normalized RG ratio of outlined regions, and finally the location,

imaging features, and predicted diagnosis of HRME sites. All acquired images are also saved, along with versions including overlays of the heat map, outlined regions, and HRME site locations, and binary masks of the outlined regions. This record keeping makes the MMIS results fully reproducible.

## 2.3 Multimodal Imaging Algorithms

### 2.3.1 Widefield image processing

**Focus metric.** The focus metric displayed during widefield imaging equals the sum of the image gradients in a circular ROI concentric to and with half the radius of full FOV. The image gradient is determined by Sobel filtering the luminance plane from HSL color space in the  $x$  and  $y$  directions and calculating the magnitude at each pixel. The center of the ROI is determined by thresholding the luminance at 15 (measured on a 0 to 255 scale) to isolate the FOV, then locating the centroid of the resulting binary image.

The focus metric only considers a portion of the full FOV because the depth of the oral cavity may be greater than the instrument's depth of field. Widefield images are typically acquired with the lesion centered, so if the circular ROI is in focus, the lesion likely is also.

**Automated normalized red to green ratio calculation.** The normalized RG ratio was calculated as previously described.<sup>15,16</sup> Briefly, the normalized RG ratio of a group of pixels is equal to their mean RG ratio divided by the mean RG ratio of a region of normal mucosa. The RG ratio of a pixel equals its red intensity divided by its green intensity. Normalization accounts for variation in illumination conditions and interpatient variability in the autofluorescence of normal tissue.

The region of normal mucosa was defined as the  $65 \times 65$  pixel square of mucosa with the lowest RG ratio. To identify this square, the RG ratio is calculated at each pixel within the mucosa, as outlined by the user. The resulting RG ratio image is filtered with a  $65 \times 65$  averaging kernel. The pixel with the lowest value after filtering is the center of the square, and its value is the RG ratio of the normal region.

**Heat map generation.** To generate the AF heat map, the normalized RG ratio of each pixel within the mucosa is calculated and blurred with a  $9 \times 9$  Gaussian filter with  $SD = 3$ . Pixels in the resultant image with a value  $>1.40$  ("neoplastic threshold") are assigned one of 64 colors from a colormap that transitions from black to red to yellow to white. The color white is assigned to all pixels with a value  $>2.20$  ("saturation threshold"). The other colors are assigned to pixels at 63 equally spaced intervals between 1.40 and 2.20. Pixels with a value  $<1.40$  are not assigned a color and are not overlaid by the heat map. These thresholds were selected based on data acquired in previous studies.<sup>14,16</sup>

**Widefield image registration algorithm.** A multiscale algorithm that aligns pairs of widefield images—each consisting of one WL image and one AF image—was developed with a training set of 40 image pairs and assessed with a test set of 28 image pairs. The image pairs were previously acquired from patients with oral lesions at MD Anderson Cancer Center (MDACC) and The University of Texas School of Dentistry (UTSD), both in Houston, Texas, as part of Institutional Review Board (IRB)-approved protocols.

The algorithm is based on mutual information (MI), a metric often used for multimodality medical image registration.<sup>17,18</sup> The MI is equal to  $h_{AF} + h_{WL} - h_{AF,WL}$ , where  $h_I$  is the entropy of image  $I$  and  $h_{I,J}$  is the joint entropy of images  $I$  and  $J$ . For an eight-bit image  $I$ ,  $h_I = -\sum_{i=0}^{255} p_I(i) \log p_I(i)$ , where  $p_I(i)$  is the probability that a pixel in  $I$  has intensity  $i$ . For two equally sized eight-bit images  $I$  and  $J$ ,  $h_{I,J} = -\sum_{i=0}^{255} \sum_{j=0}^{255} p_{I,J}(i,j) \log p_{I,J}(i,j)$ , where  $p_{I,J}(i,j)$  is the probability that the intensity at a randomly selected pixel in image  $I$  is  $i$  and the intensity at the same pixel in image  $J$  is  $j$ . Conceptually, better aligned images have higher  $p_{I,J}(i,j)$  and thus higher MI.

First, the widefield images are preprocessed, consisting of conversion to grayscale, cropping to the square inscribed in the circular FOV, and blurring with a  $5 \times 5$  Gaussian lowpass filter with  $SD = 2$ . Next, a 3-D plot of MI for  $x$  and  $y$  translations ranging from  $-100$  to  $+100$  pixels with a step size of four pixels is calculated. A preliminary, low-resolution translation is determined by selecting one of the local maxima of the MI plot based on its MI value and the gradient of its neighboring pixels. Specifically, the Sobel gradient magnitude of the MI plot is blurred with a disk filter modified to have a center value of zero. The modification is necessary so that only neighboring pixels contribute. The local maximum with the highest blurred gradient is identified, and all local maxima with a blurred gradient at least 95% of the maximal blurred gradient are retained. The retained local maximum with the highest MI is selected as the preliminary translation.

To improve precision, a 3-D plot of MI for  $x$  and  $y$  translations ranging from  $-4$  to  $+4$  pixels with a step size of 1 centered at the preliminary translation is calculated. The translation with the maximal MI is chosen as the final, high-resolution translation. The algorithm was written in Python 3.6 and compiled as an executable. The runtime is  $\sim 10$  s; parallel processing is used during MI calculations to improve efficiency.

To test the algorithm, six pairs of control points were manually selected for each pair of images in the 28 image pair test set. Then, the root mean square distance (RMSD) between the location of the control points in the WL image and the AF image was calculated for each image pair before and after registration.

### 2.3.2 High-resolution microendoscopy image processing

HRME images were processed with a previously described algorithm.<sup>13,16</sup> Briefly, the algorithm segments individual nuclei and classifies them as abnormal or normal based on their area and eccentricity. Segmentation is performed by first removing large areas of debris, enhancing image contrast and converting the result to a binary image, separating clustered nuclei with watershed segmentation, and finally removing smaller debris. The remaining objects are considered nuclei, and are classified as abnormal if their area is above  $200 \mu\text{m}^2$ , or if their area is above  $170.8 \mu\text{m}^2$  and their eccentricity is above 0.705. Otherwise, they are classified as normal. This classification is then used to calculate the number of abnormal nuclei/ $\text{mm}^2$ .

### 2.3.3 Multimodal imaging classification algorithm

In a previous study, sites in patients with oral lesions were imaged with AF and HRME.<sup>16</sup> The normalized RG ratio and number of abnormal nuclei/ $\text{mm}^2$  were calculated retrospectively, and a linear classifier was trained using linear

discriminant analysis. Since that study was published, newer versions of the AF instrumentation (introduced here) and the HRME instrumentation<sup>13</sup> were developed. We conducted a similar study using the updated instrumentation (unpublished) and updated the linear threshold. This updated threshold was used prospectively in the MMIS, which classifies sites as “dysplasia or cancer” if  $(47.1 \times \text{normalized RG ratio}) + (\text{number of abnormal Nuclei/mm}^2) \geq 273.7$ , and as “benign” otherwise.

### 2.4 Patient Imaging

Imaging of human subjects with the MMIS was performed in accordance with IRB-approved protocols at MDACC and the UTSD, both in Houston, Texas. Written informed consent was obtained from all subjects prior to imaging.

Patients with at least one oral lesion were identified and assessed by an experienced head and neck surgeon (A.G.) at MDACC or an experienced oral pathologist (N.V.) at UTSD. The clinician then decided on their clinical management (surgery, biopsy, or no biopsy) per standard of care. Next, the patients were evaluated with the MMIS, as described in the previous section. Finally, the clinicians performed biopsies or surgery per standard of care, and also had the option to biopsy patients based on MMIS results. Biopsies were 4 mm in diameter and were processed with standard MDACC or UTSD procedures with subsequent review by a study pathologist (M.W. at MDACC and N.V. at UTSD).

## 3 Results

### 3.1 Multimodal Imaging Algorithms

Figure 4 illustrates a successfully registered widefield image pair [Figs. 4(a) and 4(b)] from the test set. The red circles

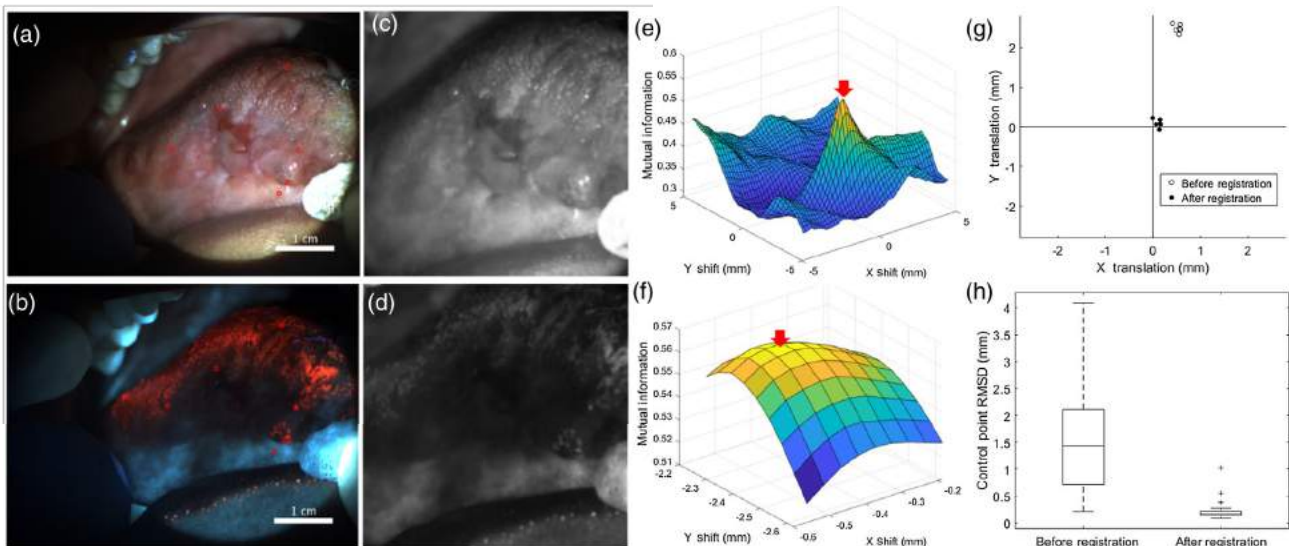
are centered at the locations of the six control points. The images were preprocessed [Figs. 4(c) and 4(d)], and MI versus  $xy$  translation was plotted [Fig. 4(e)]. The local maximum (arrow) was selected as the preliminary translation. In this example, the gradient information did not affect the preliminary translation because the local maxima with the highest MI also had the highest gradient. The translation was refined at a higher resolution [Fig. 4(f), arrow]. Figure 4(g) shows the distance between the locations of the control points in the WL and AF images before registration (open circles, RMSD 2.54 mm) and after registration (filled circles, RMSD 0.18 mm). Boxplots of the control point RMSD for all 28 test set image pairs before and after registration are shown in Fig. 4(h). Registration decreased the median RMSD from 1.44 to 0.17 mm.

Figure 5 shows an image pair [Figs. 5(a) and 5(b)] for which the gradient information was necessary for successful registration. After preprocessing [Figs. 5(c) and 5(d)], the local maxima [Fig. 5(e), filled arrow] was selected as the initial translation and was refined to a final translation [Fig. 5(f), arrow] at a high resolution. Registration decreased the control point RMSD from 0.96 [Fig. 5(g), open circles] to 0.20 mm [Fig. 5(g), filled circles]. Unlike the first example, the local maximum with the highest MI [Fig. 5(e), open arrow] was not selected as the initial translation. The gradient information was needed to select the correct local maxima.

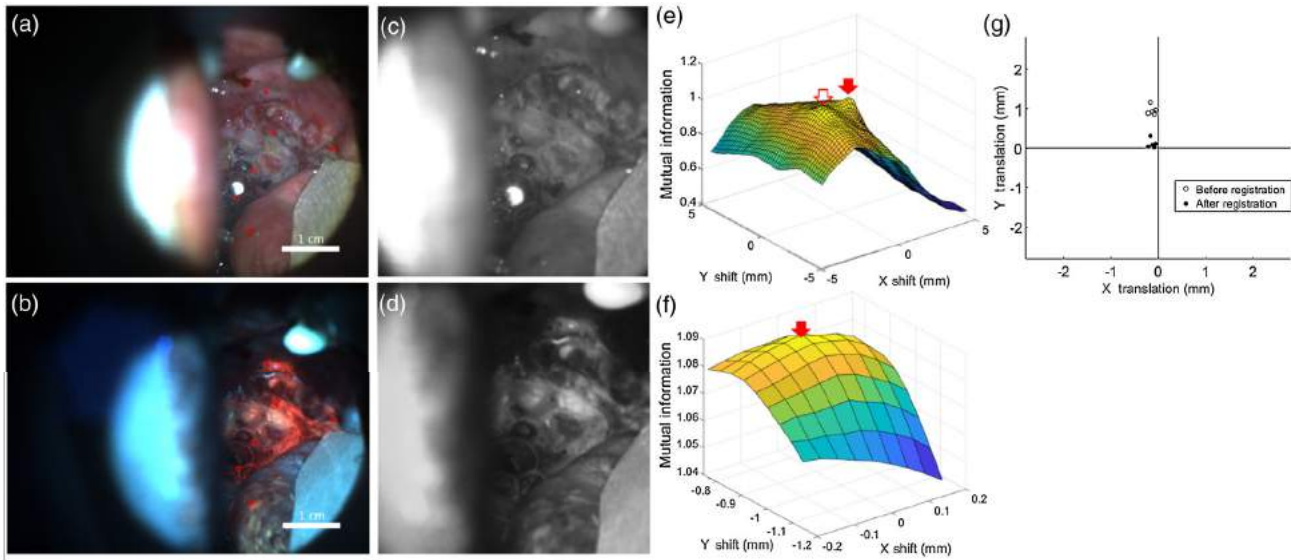
### 3.2 Patient Imaging

Selected, representative patients from ongoing studies of the MMIS are presented to illustrate its potential clinical benefit.

Figure 6 shows the use of the MMIS on an MDACC patient with a right ventral tongue lesion. The head and neck surgeon suspected cancer, and the patient was evaluated with the MMIS immediately prior to a scheduled surgical resection. The



**Fig. 4** Image registration algorithm that aligns pairs of widefield WL and AF images. (a) and (b) Example of a WL and AF image pair. The centers of the red circles represent control points. (c) and (d) WL and AF images after preprocessing. (e) MI versus  $x$  and  $y$  translation at a low resolution. The preliminary translation is indicated by the arrow. Units have been converted from pixels to millimeters. (f) MI versus  $x$  and  $y$  translation at a high resolution, centered at the preliminary translation. The final translation is indicated by the arrow. Units have been converted from pixels to millimeters. (g) Distance between corresponding control points before (open circles) and after (filled circles) registration, indicating successful registration. (h) Boxplots of the RMSD between corresponding control points for the 28 test set image pairs before and after registration.



**Fig. 5** Example of image registration where the gradient information was necessary for successful registration. (a) and (b) WL and AF image pair. The centers of the red circles represent control points. (c) and (d) WL and AF images after preprocessing. (e) Mutual information versus x and y translation at a low resolution. The preliminary translation is indicated by the filled arrow. Note that this translation was not the local maxima with the highest MI (unfilled arrow). It was selected based on its large gradient. Units have been converted from pixels to millimeters. (f) Mutual information versus x and y translation at a high resolution, centered at the preliminary translation. The final translation is indicated by the arrow. Units have been converted from pixels to millimeters. (g) Distance between corresponding control points before (open circles) and after (filled circles) registration, indicating successful registration.

widefield images shown in Figs. 6(a) (WL) and 6(b) (AF) were acquired, and the surgeon outlined a clinically suspicious region (green outline). An additional suspicious region based on the heat map [Figs. 6(c) and 6(d), red outline] was also identified. The regions overlapped almost exactly. Three HRME images were saved; sites 1 and 2 were within both regions, and site 3 was along the border [Figs. 6(e) and 6(f)]. Sites 1 and 2 had abnormal appearing nuclei and were diagnosed by the MMIS as “dysplasia or cancer” (site 1: normalized RG ratio of 2.18 and number of abnormal nuclei/mm<sup>2</sup> of 277; site 2: normalized RG ratio of 2.52 and number of abnormal nuclei/mm<sup>2</sup> of 223). Site 3, with a normalized RG ratio of 1.33 and number of abnormal nuclei/mm<sup>2</sup> of 206, was identified by the MMIS as “benign,” although it was very close to the decision boundary. All three sites were within the surgically resected specimen. Sites 1 and 2 were pathologically diagnosed as squamous cell carcinoma, and site 3 was pathologically diagnosed as mild dysplasia. Therefore, in this patient, the MMIS results, surgeon’s clinical impression, and pathologic diagnoses were all in close accord.

Figure 7 shows the use of the MMIS on a patient with a history of oral squamous cell carcinoma presenting to the MDACC clinic with leukoplakias on the right and left ventral tongue. Clinically, the lesions were not of sufficient concern to warrant a biopsy based on standard of care. The MMIS evaluated both the right ventral tongue (results not shown) and the left ventral tongue. The widefield images of the left ventral tongue are shown in Figs. 7(a) (WL) and 7(b) (AF). The surgeon outlined two regions based on clinical appearance (green outlines) and a third region based on the heat map [Figs. 7(c) and 7(d), red outline]. The HRME probe revealed abnormal appearing nuclei throughout the regions. Two HRME images were saved [Figs. 7(g) and 7(h)] and located [Figs. 7(e) and 7(f)]. Analysis

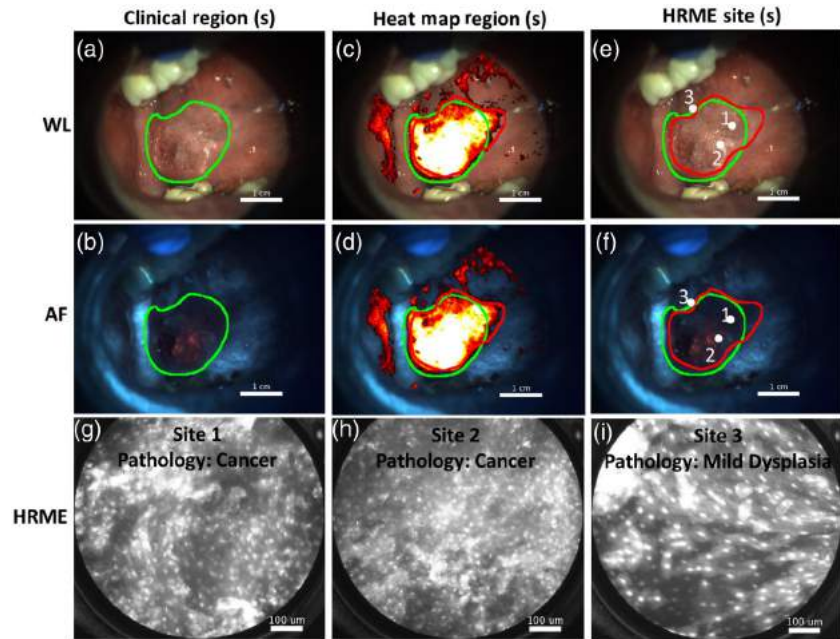
revealed an elevated number of abnormal nuclei/mm<sup>2</sup> of 255 and 319 at sites 1 and 2, respectively, both of which were diagnosed by the MMIS as “dysplasia or cancer.” Based on the MMIS results, the surgeon biopsied the mucosa at site 2, which revealed moderate to focal severe dysplasia. Therefore, in this patient, the MMIS identified focal severe dysplasia that would not have been detected by standard of care.

Figure 8 shows the use of the MMIS on a patient presenting to UTSD clinic with a left ventral tongue leukoplakia. The oral pathologist felt the lesion was of sufficient concern to warrant a biopsy based on standard of care. The MMIS acquired the widefield images shown in Figs. 8(a) (WL) and 8(b) (AF) (no clinically suspicious regions were outlined). A heat map was generated and adjusted to display fewer pixels [Figs. 8(c) and 8(d)]. A heat map region was outlined [Figs. 8(c) and 8(d), red outline]. Four HRME images within the region [Figs. 8(g)–8(j)], all with normal appearing nuclei, were saved and located [Figs. 8(e) and 8(f), white circles]. Although the normalized RG ratios at all four sites were elevated (range: 1.90 to 2.08), their low number of abnormal nuclei/mm<sup>2</sup> (range: 62 to 92) downgraded their MMIS diagnoses to “benign.” The biopsy included all four sites and revealed benign lichenoid mucositis with reactive atypia. Therefore, in this patient, the MMIS could have prevented an unnecessary biopsy.

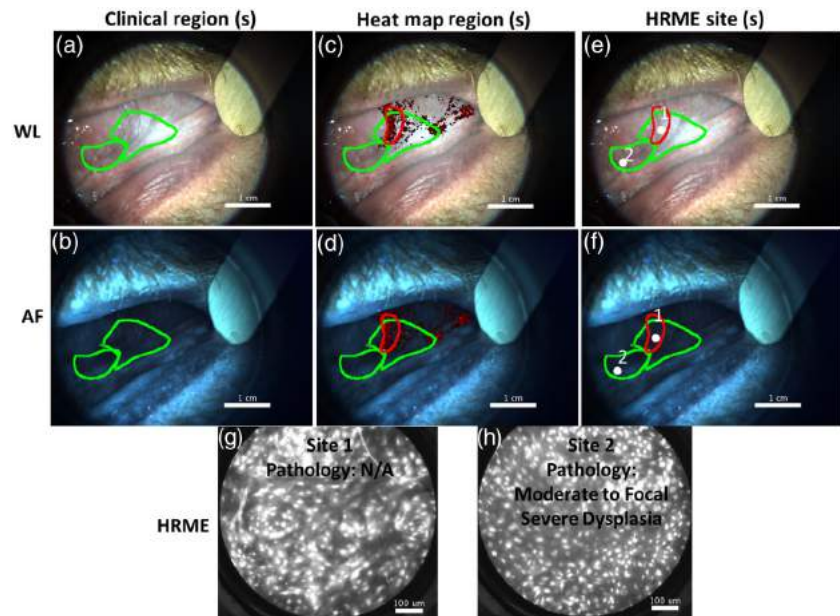
## 4 Discussion

In this paper, we developed an MMIS to evaluate oral lesions. WL and AF images are used to macroscopically identify suspicious regions, which are explored at higher resolution with the HRME. The MMIS integrates image acquisition, display, and real-time image analysis from all three modalities to predict pathologic diagnosis at imaged sites. We then present representative examples from ongoing clinical studies that illustrate the

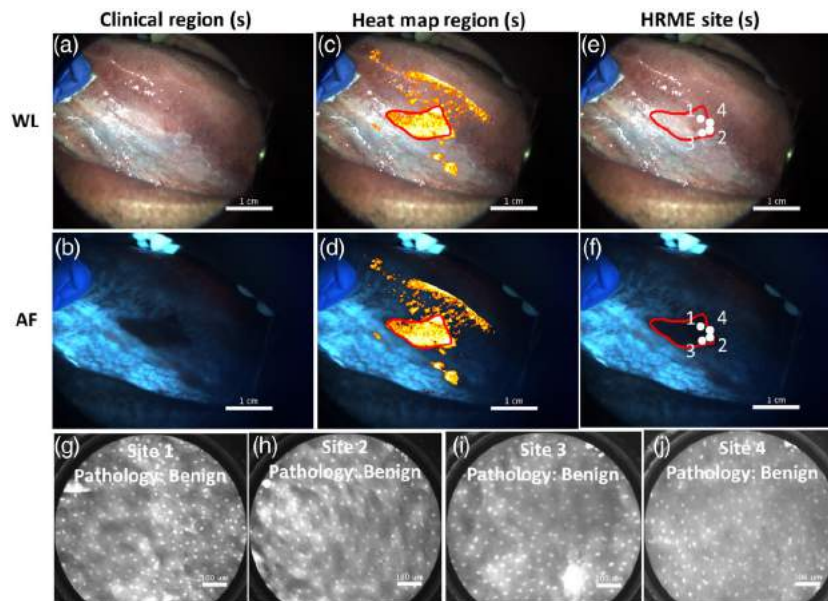




**Fig. 6** Use of MMIS on a patient immediately prior to surgical resection. (a) and (b) WL and AF image of lesion suspicious for cancer and scheduled for surgical resection. A clinical region, identified by a head and neck surgeon, was outlined (green). (c) and (d) WL and AF image including heat map overlay. An additional suspicious region based on the heat map was outlined (red). (e) and (f) WL and AF images, with HRME sites indicated by white dots. (g), (h), and (i) HRME images acquired from the sites indicated in panels (e) and (f), with corresponding pathologic diagnosis.



**Fig. 7** Example of MMIS identifying focal severe dysplasia that would not otherwise have been identified. (a) and (b) WL and AF images of lesion. Two clinical regions were outlined (green). A biopsy was not clinically indicated. (c) and (d) WL and AF images including heat map overlay. An additional suspicious region based on the heat map was outlined (red). (e) and (f) WL and AF images, with HRME sites indicated (white dots). (g) and (h) HRME images acquired from the sites indicated in panels (e) and (f). A biopsy was acquired at site 2 based on the MMIS results and revealed moderate-to-focal severe dysplasia.



**Fig. 8** Example where MMIS could potentially have prevented an unnecessary biopsy. (a) and (b) WL and AF images of lesion. A biopsy was clinically indicated, although no clinically suspicious regions were outlined. (c) and (d) WL and AF images including heat map overlay. An additional suspicious region based on the heat map was outlined (red). (e) and (f) WL and AF images, with four HRME sites indicated (white dots). (g), (h), (i), and (j) HRME images acquired from the four sites indicated in panels (e) and (f). The biopsy included all four sites and was diagnosed as benign.

potential clinical role of the MMIS. In the first example, the MMIS, expert clinical impression, and pathology were all in close accord. In the second example, the MMIS resulted in a biopsy that would not have been performed otherwise, that revealed focal severe dysplasia. Although the biopsy was acquired at site 2, which was selected based on both the MMIS results and the clinical impression, a biopsy at site 1, which was identified based on the heat map, could also have revealed dysplasia. In the third example, the MMIS diagnosed a lesion as benign; it was biopsied based on clinical suspicion for dysplasia and confirmed as benign. Together, these examples demonstrate how the MMIS could improve patient care. The MMIS can provide automated image interpretation that may be helpful for less-experienced clinicians; it can improve the ability to recognize areas of dysplasia, and it can help avoid unnecessary biopsies.

The automated, integrated image analysis of the MMIS overcomes a number of challenges associated with optical imaging. It is challenging to correlate the appearance of the mucosa under AF imaging with its appearance to the naked eye. To address this, we developed a registration algorithm that aligns the WL and AF images with a simple translation. With registration, the clinician can use the WL image instead of the AF image to outline the mucosa, outline clinically suspicious regions, outline suspicious regions with the heat map, and locate HRME sites. Registration accuracy could be improved by assuming a more complex transformation, but the median misalignment after registration corresponds to only a few hundred microns [Fig. 4(h)], a smaller distance than the precision with which HRME probes are manually placed.

The ideal diagnostic adjunct requires minimal training to use and interpret. This criterion is particularly important if the adjunct is to be used by nonspecialists. Toward this goal, the MMIS uses validated, automated algorithms to classify

imaged sites as benign or moderate dysplasia-cancer. However, clinical judgment does still play a role. The clinician identifies the lesions that are to be evaluated by the MMIS. Clinical judgment is also used to outline the clinically suspicious regions and the heat map regions. Previous studies have shown that normal gingival and buccal mucosa can have increased red fluorescence.<sup>7,14–16</sup> These regions can be falsely highlighted by a heat map with a single threshold. Imaging artifacts such as light reflected off the metal instruments used to expose the lesion and shadows can also alter the heat map. Trained users can recognize these artifacts and outline regions accordingly. For patients with larger lesions, the mucosa initially highlighted by the heat map may be too large to fully explore with the HRME probe. The user can decrease the size of the heat map until it is feasible, but this could introduce interoperator variability.

In summary, the MMIS is a minimally invasive, point-of-care diagnostic adjunct with the potential to help clinicians decide whether an oral lesion should be biopsied and to select a biopsy site. These improvements could decrease the global oral cancer burden. Future work will focus on continuing to image OPL patients so that the MMIS's ability to evaluate oral lesions can be assessed systematically. The additional data could also be used to compare the ability of clinical impression to autofluorescence in identifying sites with altered nuclear morphology and improve the diagnostic algorithms. Improvements to the instrumentation, such as utilizing ring light AF illumination to decrease shadows, and additional software features, can also be made.

### Disclosures

R. Richards-Kortum, A. M. Gillenwater, and R. A. Schwarz are recipients of licensing fees for intellectual property licensed from the University of Texas at Austin by Remicalm LLC.

All other authors disclosed no potential conflicts of interest relevant to this publication.

### Acknowledgments

This work was supported by the National Institutes of Health Grant Nos. RO1 CA103830 (to R. Richards-Kortum); RO1 CA185207 (to R. Richards-Kortum); RO1 DE024392 (to N. Vigneswaran); F30CA213922 (to E. Yang), and by the Cancer Prevention and Research Institute of Texas (CPRIT) Grant No. RP100932 (to R. Richards-Kortum).

### References

1. J. Ferlay et al., "GLOBOCAN 2012 v1.0, cancer incidence and mortality worldwide: IARC CancerBase No. 11," International Agency for Research on Cancer, Lyon, France, 2013, <http://globocan.iarc.fr/Default.aspx> (13 November 2016).
2. M. W. Lingen et al., "Evidence-based clinical practice guideline for the evaluation of potentially malignant disorders in the oral cavity: a report of the American Dental Association," *J. Am. Dent. Assoc.* **148**, 712–727 (2017).
3. S. Warnakulasuriya et al., "Oral epithelial dysplasia classification systems: predictive value, utility, weaknesses and scope for improvement," *J. Oral Pathol. Med.* **37**, 127–133 (2008).
4. J.-J. Lee et al., "Factors associated with underdiagnosis from incisional biopsy of oral leukoplakic lesions," *Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endodontology*. **104**, 217–225 (2007).
5. R. Macey et al., "Diagnostic tests for oral cancer and potentially malignant disorders in patients presenting with clinically evident lesions," in *Cochrane Database of Systematic Reviews*, R. Macey, Ed., p. 4, John Wiley and Sons, Ltd., Hoboken, New Jersey (2015).
6. I. Pavlova et al., "Understanding the biological basis of autofluorescence imaging for oral cancer detection: high-resolution fluorescence microscopy in viable tissue," *Clin. Cancer Res.* **14**, 2396–2404 (2008).
7. D. Roblyer et al., "Objective detection and delineation of oral neoplasia using autofluorescence imaging," *Cancer Prev. Res.* **2**, 423–431 (2009).
8. I. Pavlova et al., "Fluorescence spectroscopy of oral tissue: Monte Carlo modeling with site-specific tissue properties," *J. Biomed. Opt.* **14**, 014009 (2009).
9. J. E. Bouquot, "Common oral lesions found during a mass screening examination," *J. Am. Dent. Assoc.* **112**, 50–57 (1986).
10. M. Kovac-Kavcic and U. Skaleric, "The prevalence of oral mucosal lesions in a population in Ljubljana, Slovenia," *J. Oral Pathol. Med.* **29**, 331–335 (2000).
11. A. L. Mathew et al., "The prevalence of oral mucosal lesions in patients visiting a dental school in Southern India," *Indian J. Dent. Res.* **19**, 99–103 (2008).
12. T. J. Muldoon et al., "Subcellular-resolution molecular imaging within living tissue by fiber microendoscopy," *Opt. Express* **15**, 16413–16423 (2007).
13. T. Quang et al., "A tablet-interfaced high-resolution microendoscope with automated image interpretation for real-time evaluation of esophageal squamous cell neoplasia," *Gastrointest. Endosc.* **84**, 834–841 (2016).
14. M. C. Pierce et al., "Accuracy of in vivo multimodal optical imaging for detection of oral neoplasia," *Cancer Prev. Res.* **5**, 801–809 (2012).
15. E. C. Yang et al., "In vivo multimodal optical imaging: improved detection of oral dysplasia in low-risk oral mucosal lesions," *Cancer Prev. Res.* **11**(8), 465–476 (2018).
16. T. Quang et al., "Prospective evaluation of multimodal optical imaging with automated image analysis to detect oral neoplasia in vivo," *Cancer Prev. Res.* **10**, 563–570 (2017).
17. F. Maes et al., "Multimodality image registration by maximization of mutual information," *IEEE Trans. Med. Imaging* **16**, 187–198 (1997).
18. J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Mutual-information-based registration of medical images: a survey," *IEEE Trans. Med. Imaging* **22**, 986–1004 (2003).

Biographies of the authors are not available.