

# Journal of Biomedical Optics

BiomedicalOptics.SPIEDigitalLibrary.org

## Red fluorescence imaging for dental plaque detection and quantification: pilot study

Zhao Liu  
Juliana Gomez  
Soniya Khan  
Debbie Peru  
Roger Ellwood

**SPIE.**

Zhao Liu, Juliana Gomez, Soniya Khan, Debbie Peru, Roger Ellwood, "Red fluorescence imaging for dental plaque detection and quantification: pilot study," *J. Biomed. Opt.* **22**(9), 096008 (2017), doi: 10.1117/1.JBO.22.9.096008.

# Red fluorescence imaging for dental plaque detection and quantification: pilot study

Zhao Liu,<sup>a,\*</sup> Juliana Gomez,<sup>b</sup> Soniya Khan,<sup>b</sup> Debbie Peru,<sup>b</sup> and Roger Ellwood<sup>a</sup>

<sup>a</sup>University of Manchester, Colgate Palmolive Dental Health Unit, Williams House, Lloyd Street North, Manchester, United Kingdom

<sup>b</sup>Colgate Palmolive Technology Center, Piscataway, New Jersey, United States

**Abstract.** The red fluorescence of dental plaque originating from porphyrins in oral bacteria may allow visualization, detection, and scoring of plaque without disclosing agents. Two studies were conducted. The first included 24 healthy participants who abstained from oral hygiene for 24 h. Dental plaque was collected from tooth surfaces, and a 10% solution was prepared. These were scanned by a molecular spectrometer to identify the optimum excitation and emission wavelengths of plaque for developing a red fluorescence imaging system. Fourteen healthy subjects completed the second study. After a washout period (1 week), participants had a prophylaxis at baseline and abstained from oral hygiene during the study. They were monitored using the fluorescence imaging system at baseline, 24 h, and 48 h. A dentist clinically assessed plaque after disclosing and on red fluorescence images. Three descriptors were extracted from images and a RUSBoost classifier derived computer fluorescence scores through cross-validation. Red fluorescence plaque levels increased during the 48-h accumulation. Plaque progression was identified by dentist assessment and computer analysis, presenting significant differences between visits at tooth and subject levels ( $p < 0.05$ ). Moderate correlations showed between clinical plaque and red fluorescence plaque ( $r = 0.62$  dentist,  $r = 0.55$  computer). The best agreement was observed when disclosing plaque threshold at level 2, for both dentist evaluation (sensitivity 71.1%, specificity 67.7%, accuracy 70.2%) and computer classification (sensitivity 68.4%, specificity 62.9%, accuracy 67.1%). Given the correlation with clinical diagnosis, red fluorescence imaging shows its potential for providing an objective and promising method for proper oral hygiene assessment. © 2017 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: [10.1117/1.JBO.22.9.096008](https://doi.org/10.1117/1.JBO.22.9.096008)]

Keywords: dental plaque; autofluorescence imaging; oral hygiene; RUSBoost classifier; computer analysis.

Paper 170374RR received Jun. 8, 2017; accepted for publication Aug. 30, 2017; published online Sep. 18, 2017.

## 1 Introduction

The accumulation of dental plaque may cause tooth decay and gum diseases as a result of the accumulation of bacteria.<sup>1</sup> Therefore, the removal of dental plaque is crucial for maintaining oral health.

Human examination still plays a dominant role in the clinical identification of dental plaque by using disclosing solutions combined with index scales based on attributes such as the plaque area and its thickness.<sup>2</sup> However, human visual assessments are subjective and often inconsistent, which generates difficulty in comparing results between baselines and follow up diagnosis. The use of disclosing agents may stain soft tissue, which could decrease specificity and cause esthetic concerns. An objective and noninvasive method to detect and quantify dental plaque, especially at the early stage of its formation, would be of value in dentistry.

Autofluorescence imaging has been used for the detection of tooth defects, such as differences in mineralization, due to the intrinsic fluorescence characteristics of teeth when irradiating with blue light.<sup>3</sup> It showed diagnostic values to support the clinical diagnosis of caries lesions and fluorosis.<sup>4-6</sup> Red fluorescence, generally believed to originate from porphyrins produced by bacterial metabolites in the anaerobic plaque,<sup>7,8</sup> has been used for dental plaque monitoring without disclosing solutions.

Two types of autofluorescence imaging systems, including Sopro camera by Acteon<sup>9</sup> and a quantitative light-induced fluorescence digital (QLF-D) system,<sup>10</sup> were reported to apply for the plaque detection in the literature. Rechmann et al.<sup>11</sup> used Sopro camera to evaluate microbial plaque level and the plaque scores output by the Soprocare assessment tool demonstrated consistent with the clinical plaque indices. However, so far, there is no published clinical study to monitor the plaque progression using Sopro camera. It is therefore unclear about the sensitivity of Sopro camera for quantifying the plaque changes. On the other hand, two recent clinical studies<sup>10,12</sup> employed the QLF-D system to compare the red fluorescence plaque with the clinical plaque and showed moderate correlations between them. But these works only included anterior teeth. Neither study reported a prophylaxis step. It is ambiguous whether the identified red fluorescence signal (especially  $\leq 48$  h) is primarily attributed from the plaque accumulated during the experimental period or built up for a longer time.

There are three primary objectives of this cross-sectional study: (1) to identify appropriate excitation and emission wavelengths of dental plaque for red autofluorescence and inform the construction of a fluorescence imaging system that allows full access into the human mouth; (2) to monitor the development of dental plaque using the proposed system and compare it to clinical scoring; and (3) to introduce an automatic plaque detection

\*Address all correspondence to: Zhao Liu, E-mail: [zhao.liu-2@manchester.ac.uk](mailto:zhao.liu-2@manchester.ac.uk)

and quantification method for understanding the feasibility of developing an objective platform for plaque assessment.

## 2 Materials and Methods

### 2.1 Plaque Collection Study

A plaque collection study was performed at the Technology Center Colgate-Palmolive Piscataway, New Jersey, in January 2016 to understand the autofluorescence characteristics of dental plaque. The study followed the instructions based on the Declaration of Helsinki (2008) and with the U.S. Code of Federal Regulations governing the protection of human subjects (21 CFR 50) and the institutional review boards (21 CFR 56). It was approved by Clinical Research Institutional Review Board, Concordia, New Jersey, under the registration number CRO-2015-1110-PLACOLL-JG.

Volunteers between 18 and 65 years old, having a minimum of 20 natural uncrowned teeth, and in good general health, were eligible to participate. They agreed to take part in this study by reading and signing an informed consent form and completed an oral hard and soft oral tissue exam, together with a general health questionnaire form.

Participants were excluded with advanced periodontal disease (purulent exudate, tooth mobility, and/or extensive loss of periodontal attachment or alveolar bone) and spontaneous bleeding of the gums. Immune compromised individuals (HIV, AIDS, immunosuppressive drug therapy), or subjects taking medications that may affect the oral flora (e.g., use of antibiotics 1 month prior to or during the study), were also excluded from the study. Further exclusion criteria were five or more carious lesions requiring immediate restorative treatment, tumor of the soft or hard tissues of the oral cavity, abnormal salivary function, premedication required for dental procedures, pregnant women, or women who were breastfeeding.

A total of 24 subjects were enrolled in the study. Participants abstained from oral hygiene at least for 24 h. During the experiment, dental plaque was collected from tooth surfaces by a certified hygienist and stored in a preweighted cylindrical test tube. Test tubes with plaque were then weighted again to determine the plaque weight. For analysis, 90% phosphate-buffered solution was added to each test tube to create a 10% plaque sample, and the tubes were vortexed for 1 min. The solution was then dispensed into a 96-well microplate and scanned by the

Molecular Device SpectraMax M5 Spectrometer (Sunnyvale, California), first in excitation mode, and then in emission mode, to determine the optimal wavelength for generating a maximum fluorescence intensity response. For each plaque sample, the optimum excitation wavelength was determined from the wavelength giving the strongest emission intensity at 650 nm when scanning excitation spectrum from 300 to 600 nm. The optimum emission wavelength was defined as the strongest emission response from 500 to 800 nm generated using an excitation wavelength of 405 nm. The optimal excitation and emission wavelengths for each subject were recorded and used to calculate a mean and standard deviation across all subjects to inform the construction of an autofluorescence imaging system for plaque detection.

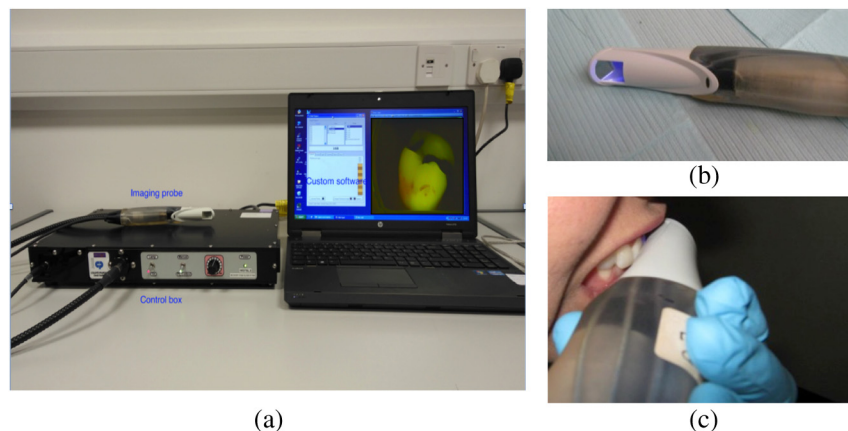
### 2.2 Red Fluorescence Imaging System

Based on the excitation and emission experiment described above (results shown in Sec. 3.1), a fluorescence imaging equipment was developed at Colgate Palmolive Dental Health Unit, University of Manchester, for the acquisition of dental plaque images of human teeth. This consisted of two parts: a handheld imaging probe and an illumination controller.

The handheld probe allowed imaging of the full mouth. It was comprised of a lipstick color camera (IK-UM51H 1/3" Toshiba) fitted with a 17.5-mm focal length lens (58206 Edmund Optics), offering an effective field of view of 13.5 mm × 16.5 mm able to cover the full range of a single incisor, canine, or molar. It is worth noting that the near-infrared filter, cutting around 650 nm, was removed from the lipstick camera to capture the red fluorescence signal in the spectrum beyond 650 nm.

The control box included a near-UV LED centered at 405 nm with power density of 37.1  $\mu\text{W}/\text{mm}^2$  (M405L2-UV, Thorlabs Inc.) to provide the excitation light for red fluorescence imaging of dental plaque. During the clinical study, participants were provided with a pair of UV-protection glasses and dentist put on UV-protection gloves across the whole image acquisition process. The emission light was filtered by a 500 nm long-pass filter (62976, Edmund Optics), which was placed in front of the camera lens in the handheld probe.

A customized mirror tip, mounted on top of the handheld probe, allowed the full access into human mouth for imaging



**Fig. 1** Red fluorescence imaging system. (a) The whole imaging system, including imaging probe, control box, and custom software. (b) Imaging probe with excitation light turned on. Blue light reflected by the disposable intraoral tip. (c) Imaging probe applied during clinical study.

both anterior and posterior teeth. An air bump was connected to the imaging probe to clear the fogging on the mirror window caused by patient breathing during the measurement. In the clinical study, this mirror tip was replaced for every patient and every visit for infection control purposes.

Both the camera and illuminator were controlled by custom-written software. Camera settings, such as gain, shutter speed, image pixel resolution, etc., were preconfigured with a study design. This study design was locked unchangeable throughout the experiments. The brightness of the near-UV light was computer controlled by changing the output of a current controller that was connected to the LEDs. The light achieved ideal brightness when the average pixel intensity of a standard gray card reached a predefined value with a small tolerance of  $\pm 0.05$  pixel intensity value. This made imaging illumination at different times comparable. During the image acquisition, the fluorescence images of human teeth were captured by moving and focusing the camera at baseline. For the follow-up visit, an edge detection overlay extracted from the baseline image was displayed on top of the live video. This overlay guided the dentists to reposition the imaging probe to capture the teeth areas as close as those from baseline, so that the plaque changes shown on the fluorescence images were comparable between visits. Figure 1 shows the red fluorescence imaging system stated above.

## 2.3 Plaque Detection Using Red Fluorescence Imaging

### 2.3.1 Study design and study population

This cross-sectional study under the protocol number CRO-2016-PLA-04-MIT-JG was conducted at the Technology Center Colgate-Palmolive Piscataway, New Jersey, in August 2016. The good clinical practice guideline, review board approval, and the inclusion and exclusion criteria were the same as those in the plaque collection study described in Sec. 2.1.

### 2.3.2 Clinical procedures

Upon completion of all enrollment requirements, participants were provided with a tube of fluoride toothpaste (Colgate Great Regular Flavor) and a Colgate adult soft bristle toothbrush to brush twice daily for 1 week prior to the start of the study.

During the experiment, participants were requested to refrain from any form of oral hygiene continuously for 48 h. Visits were scheduled at baseline, 24 h, and 48 h. At the baseline, subjects received a prophylaxis before the clinical and imaging assessments.

For each visit, six red fluorescence images were captured at the subject's buccal surfaces of the upper first molars (16, 26), the buccal surfaces of the upper laterals (12, 22), and the buccal surfaces of first lower molars (36, 46) using the described red fluorescence imaging system. All the images should be focused and the target teeth were displayed in the image center. Next, dental plaque was disclosed by gently applying a fluorescein dye (Plaque Test; Ivoclar Vivadent, Liechtenstein) on the tooth surfaces. The mouth was rinsed twice and subsequently dried so that the fluorescein dye adhered to the areas covered with plaque. Finally, a trained and certified dentist clinically assessed the disclosing dental plaque by illuminating the teeth with a blue light (Schott KL1600 LED filtered by a short-pass filter cutting at 486 nm).

In this study, clinical disclosing plaque was scored using the Turesky modification of the Quigley Hein plaque index (T-QH).<sup>3</sup> Dentist recorded "cannot access" when the disclosing plaque was visually difficult to achieve. Three T-QH scores were given for each tooth describing the plaque level at distal, buccal, and mesial sites.

### 2.3.3 Red fluorescence plaque scores

Anonymized red fluorescence images were visually assessed by a single trained and calibrated examiner using custom-scoring software. T-QH was reported adaptive for use on the fluorescence plaque images.<sup>10,12</sup> Considering the extent of fluorescence plaque accumulated within 48 h after prophylaxis, a modified T-QH, having a four point scale (0 to 3), was applied in this study by merging the original T-QH scores 3 to 5 into one plaque level. This red fluorescence plaque index was referred to RFT-QH in the following context. Similar to disclosing plaque scores, one more option, "cannot access," was also listed in the dropdown menu in the scoring software, to supplement the cases when the red fluorescence plaque assessment was not available.

During the evaluation, images were presented to the examiner randomly and three RFT-QH scores were given at distal, buccal, and mesial sites for each tooth. When finished, the dentist reviewed the recorded fluorescence scores and made changes if necessary. All the scores were thereafter exported from the software for subsequent analysis.

For analysis purpose, the tooth level index was defined as the maximum assessment among three site scores within each tooth regardless of the index "cannot access." The subject level score was the average value of the tooth level indices of each patient. Table 1 summarizes the dentist's clinical evaluation and red fluorescence plaque scores at site and tooth levels.

## 2.4 Automatic Plaque Detection and Classification on Fluorescence Images

### 2.4.1 Automatic plaque detection

Associating autofluorescence characteristics with the spectral responses of conventional RGB cameras, plaque areas present

**Table 1** Summary of dentist's clinical evaluation using disclosing solution and dentist's red fluorescence plaque scores based on fluorescence images.

T-QH/RFT-QH	Clinical scores		Dentist RFT-QH scores	
	Site level	Tooth level	Site level	Tooth level
0	173	20	464	97
1	115	42	155	86
2	338	134	73	45
3	71	48	32	24
4	9	8	—	—
5	0	0	—	—
Cannot access	0	0	32	—
Total	756	252	756	252

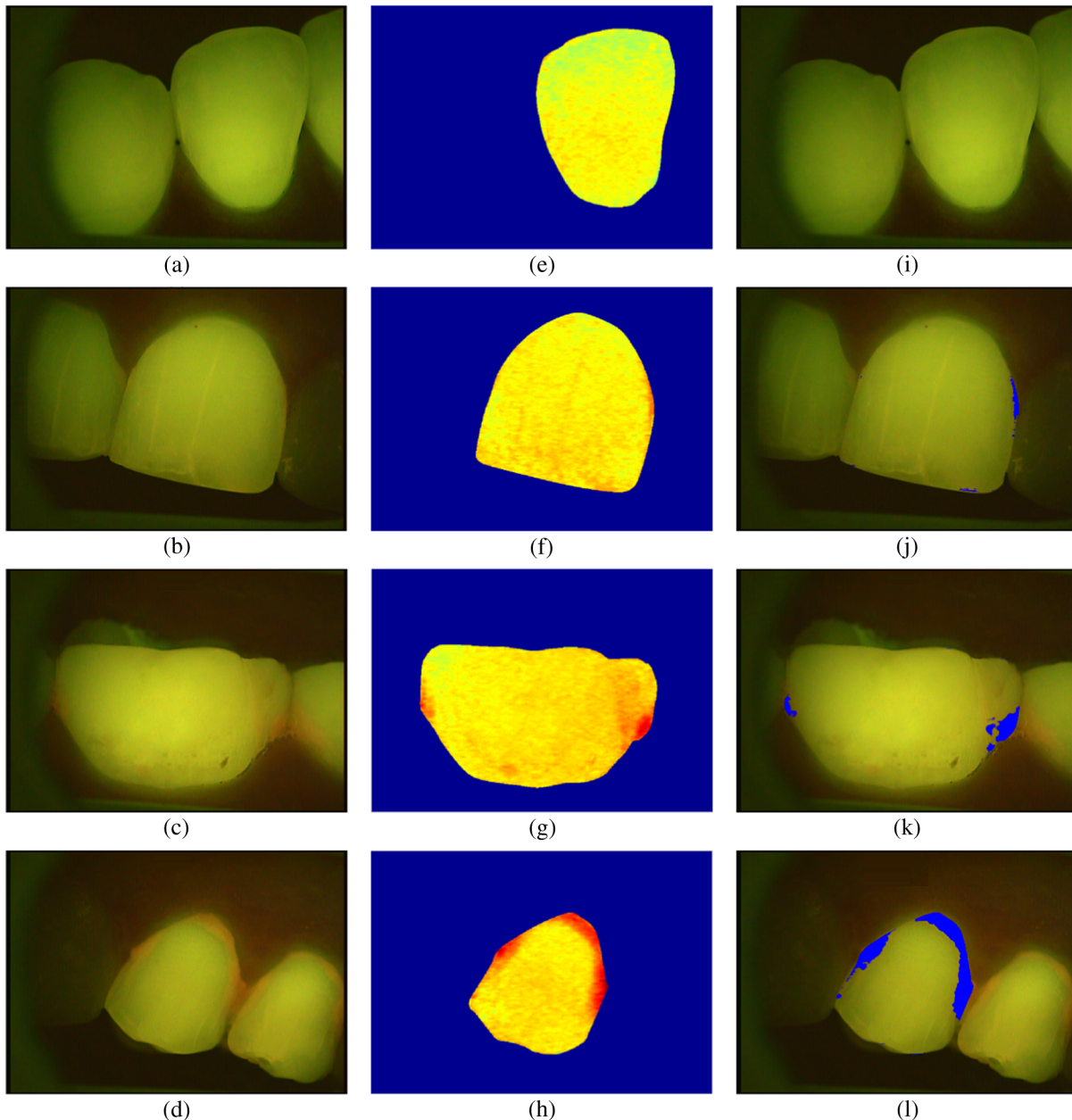
apparent higher intensity in the red channel while the healthy teeth fluoresce green due to the intrinsic property of hard dental tissues. Dental plaque distribution  $D_p$  therefore can be revealed by the quotient between two spectral bands, as defined in Eq. (2). According to the Lambertian law,<sup>13</sup> this division cancels out the shading effects due to tooth curvature:

$$D_p = \frac{I_{\text{Red}}}{I_{\text{Green}}}, \quad (2)$$

where  $I_{\text{Red}}$  and  $I_{\text{Green}}$  stand for the red and green channels of the fluorescence image, respectively.

A binary mask  $M_p$  can be identified by setting threshold in the plaque feature space  $D_p$ . Pixels having intensity above the threshold were classified as plaque and labeled 1, and the rest of the pixels were considered as healthy dental tissue marked 0. In this study, the plaque feature was regulated to an interval of 0 to 5. The threshold was set to 1.0 as a constant throughout the work, which means 20% of the intensity range in the feature space  $D_p$ .

For a selected tooth outlined by the dentist, the average  $D_p$  value of nonplaque pixels was considered as the intensity level of sound adjunct enamel. Therefore, a plaque intensity measure  $\Delta P$  can be calculated for each plaque pixel within the chosen tooth:



**Fig. 2** Red fluorescence images scored 0 to 3 by the examiner at tooth level with detected plaque areas displayed. (a)–(d) Red fluorescence images scored 0 to 3 (from top to bottom) by a certified dentist. (e)–(h) Dental plaque distribution  $D_p$  of the selected tooth, where green/yellow color refers to small  $D_p$  value and orange/red colour stands for large one. (i)–(l) Detected plaque areas (blue color) within the selected tooth overlaid on top of fluorescence images.

$$\Delta P(x) = D_{p, M_p=1}(x) - \overline{D_{p, M_p=0}}, \quad (3)$$

where  $x$  represents a plaque pixel, and  $\overline{D_{p, M_p=0}}$  is the average intensity of sound adjunct enamel in the plaque feature space.

A three-dimensional feature vector, including the average plaque intensity  $\Delta P_{ave}$ , the area percentage of plaque pixels within the whole tooth  $\delta$ , and their dot product  $\Delta P_{ave} * \delta$ , was calculated. These features were then input into an ensemble classifier as the diagnostic descriptors to estimate the plaque level on each tooth.

Figure 2 shows the red fluorescence images scored 0 to 3 by the certified dentist at tooth level with the detected plaque areas overlaid.

#### 2.4.2 Dental plaque classification

From Table 1, it should be noted that the number of red fluorescence plaque scored 0 was approximately 10 times (464/32) of those marked 3 at the site level and 5 times (97/24) at the tooth level. Hence, the experimental dataset in the present study was imbalanced. Traditional machine learning algorithms would be ineffective at dealing with this skewed data, since these methods tend to favor classifying observations as belonging to the majority class. As such, an ensemble classifier, random undersampling boost (RUSBoost),<sup>14</sup> was employed to alleviate the problem of class imbalance.

RUSBoost is a hybrid learning strategy to produce a predictive model through integrating the outputs from a finite number of weak learners (also known as base classifiers) that train for the same task. It centers around two considerations: (1) data sampling in order to balance the class distribution and (2) boosting focusing on improving the classification performance through iteratively creating an ensemble of weak hypotheses.

Random undersampling (RUS) randomly removes observations from majority class in the training data until a desired class distribution is achieved. Although there is no universally accepted rule for setting class ratios, an equal distribution of all the classes is normally considered working well for most applications. Adaptive boosting (AdaBoost),<sup>15</sup> one of the most common boosting algorithms, is an iterative process, which aims at correcting the instances misclassified by the previous learner in the subsequent one. Specifically, during one iteration in AdaBoost, a hypothesis is constructed by a base classifier. Then, the weight associating with each observation is adjusted so that weights of misclassified observations increase and those of correctly classified examples decrease. Therefore, previously mislabeled data are most likely to be classified correctly in the succeeding iteration. Upon completion, the hypotheses of testing data from all weak learners are fused via a weighted vote, and the object is assigned to the class giving the maximum weight.

In this study, decision tree derived from C4.5 algorithm<sup>16</sup> using an entropy-based splitting criterion was applied as the weak learner. AdaBoost.M2<sup>17</sup> minimizing the pseudo-loss in each iteration was the boosting method for creating an ensemble of weak hypotheses. An equal distribution of all the classes was achieved for training the weak learners using RUS approach, and the number of iteration in AdaBoost was set to 100.

#### 2.5 Experiments and Statistic Analysis

In this pilot study, a number of 25 subjects were screened, of which 20 met the inclusion criteria and 15 were enrolled

after informed consent. One participant dropped out at 24 h visit during the study. Hence, this paper reported the results based on the 14 participants who followed the entire protocol.

Excluding the scores of "cannot access," there were 756 clinical disclosing plaque scores and 724 dentist RFT-QH scores included in the experiment. A threefold cross-validation<sup>18</sup> was combined with the ensemble classifier as the training and testing strategy to derive the computer red fluorescence plaque scores at tooth level.

The performance of plaque classification was evaluated using confusion matrix. Sensitivity (true positive number/number of plaque cases), specificity (true negative number/number of healthy cases), and the total agreement (true positive number + true negative number/total number)<sup>19</sup> were also recorded. A linear weighted kappa<sup>20</sup> was performed to measure the level of agreement between dentist diagnosis and automatic classification. Differences of computer features (plaque intensity, plaque areas, and their product) between plaque levels (RFT-QH) were calculated by  $t$ -test, and those between different visits were analyzed by paired  $t$ -test.

Spearman rank correlation<sup>21</sup> was used to determine the association between disclosing plaque indices with red fluorescence RFT-QH scores and computer features. Differences of plaque scores between visits were calculated by Wilcoxon signed-rank test.

Based on the results derived from this pilot study, a *post-hoc* power analysis<sup>22</sup> was performed to outline the sample size for future clinical trials. Power analysis was conducted using the distributions of computer plaque scores for differentiating plaque level at baseline from that after 48 h plaque accumulation.

The automatic plaque detection and classification algorithms were implemented using C++ under Visual Studio 2010 (Microsoft, United States), running on a platform of Intel i7-4770 at 3.40 GHz CPU and 8GB DDR3-1600 RAM. Statistical analysis was performed using SPSS (version 23, IBM Inc., United States), and *post-hoc* power analysis was conducted using R.

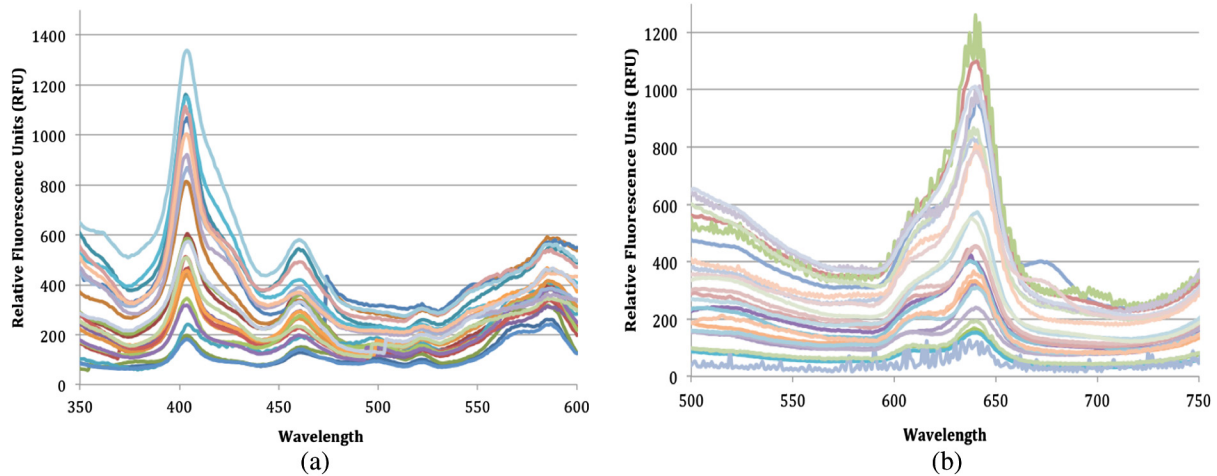
### 3 Results

#### 3.1 Excitation and Emission Spectrum of Dental Plaque

From Fig. 3, it is noted that there was little variation of the excitation and emission spectrum of dental plaque in peak wavelength among subjects. The average excitation peak of 24 subjects was  $403.79 \pm 0.59$  nm and the average emission peak was  $639.75 \pm 1.26$  nm. This was consistent to the previous findings, which claimed that porphyrins have the excitation band around 400 nm and the emission wavelength after 600 nm.<sup>7,8</sup>

#### 3.2 Automatic Classification Accuracy of Red Fluorescence Plaque

Table 2 shows the automatic classification results of red fluorescence plaque derived from the proposed methods taking the dentist RFT-QH assessment as the reference standard. The total classification accuracy was 62.7%, where the true positive rates of plaque level 0 to plaque level 3 are 67%, 50%, 68.9%, and 79.1%, respectively. A moderate agreement showed between dentist RFT-QH scores and computer RFT-QH scores with a linear weighted kappa of 0.594.



**Fig. 3** Excitation and emission spectrum of dental plaque from 24 subjects measured by spectraMax M5 spectrometer. Colors represent fluorescence characteristics of plaque from different subjects. (a) Excitation spectrum from 350 to 600 nm in excitation mode. (b) Emission spectrum from 500 to 750 nm in emission mode generated using an excitation wavelength of 405 nm.

**Table 2** Automatic classification results (number/percentage) of red fluorescence plaque at tooth level using RUSBoost classifier with threefold cross-validation. Values in the left diagonal cells represent the true-positive number and rate in each class. Values in other cells show the number and rate of misclassified cases of different plaque levels.

PC RFT-QH	Dentist RFT-QH			
	0	1	2	3
0	65 (67%)	27 (31.4%)	3 (6.7%)	0 (0%)
1	21 (21.7%)	43 (50%)	6 (13.3%)	1 (4.2%)
2	10 (10.3%)	16 (18.6%)	31 (68.9%)	4 (16.7%)
3	1 (1%)	0 (0%)	5 (11.1%)	19 (79.1%)

Figure 4 demonstrates the distributions of computer features at different plaque levels (RFT-QH) for both dentist assessment and computer classification. Plaque intensity ( $\Delta P_{ave}$ ) increased and the plaque areas ( $\delta$ ) enlarged with the increment of RFT-QH scores. Taking the significant level at  $p < 0.05$ , all the computer features resulted in significant differences between any two plaque levels within the RFT-QH scale.

### 3.3 Correlation Between Clinical Disclosing Plaque and Red Fluorescence Plaque

From Table 3, a moderate to strong correlation was identified between clinical disclosing plaque and red fluorescence plaque at subject level. Tooth level association was overall lower presenting weak to moderate correlations. Plaque areas ( $\delta$ ) and inner product ( $\Delta P_{ave} * \delta$ ) correlated well with dentist RFT-QH scores at tooth and subject levels throughout the experiment. Their correlations with clinical assessment were moderate to strong at subject level and weak to moderate at tooth level (Table 4). This observation was consistent to the results between clinical plaque scores and RFT-QH scores based on the red

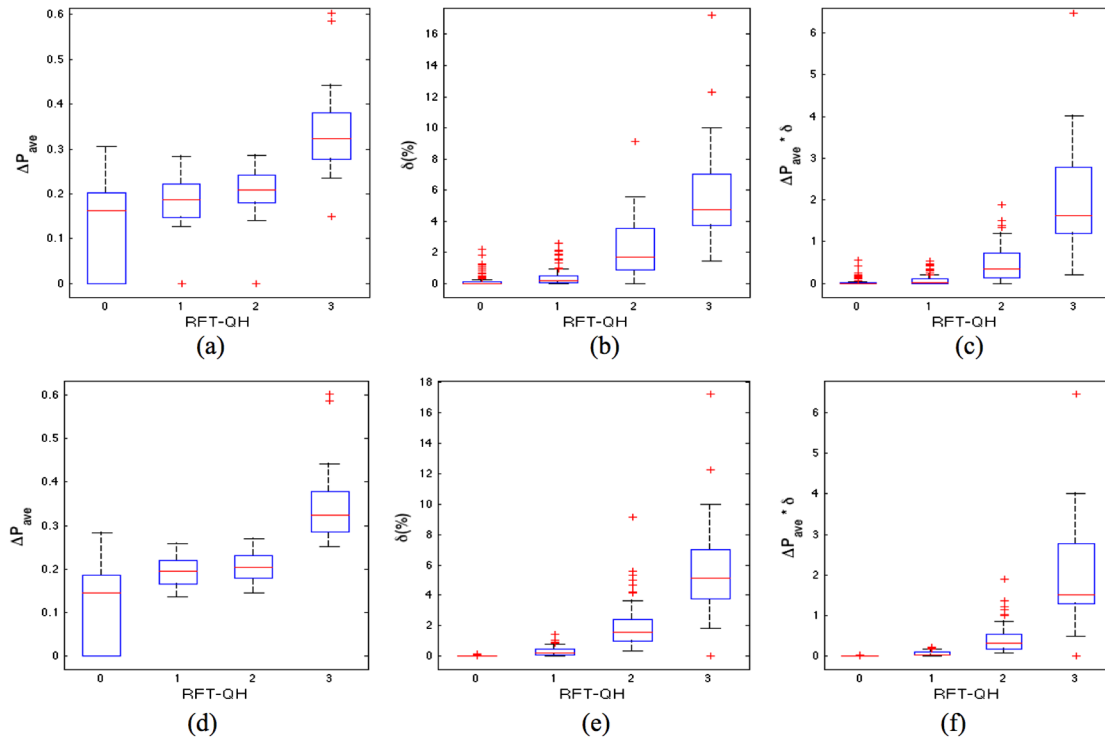
fluorescence images. The average plaque intensity ( $\Delta P_{ave}$ ) was less relevant with human visual assessment giving moderate correlations at most.

Considering red fluorescence score  $RFT - QH \geq 1$  as plaque affected, Fig. 5 demonstrates the agreement between clinical disclosing plaque indices and red fluorescence plaque scores, taking the disclosing parameters as assessment references. It is noted that, with the increase of threshold in disclosing T-QH, sensitivity of plaque detection by the red fluorescence imaging increased while the corresponding specificity decreased. The agreement between dentist disclosing scores and computer RFT-QH scores presented comparable results with that between disclosing scores and dentist RFT-QH assessment. The best agreement of plaque detection was observed, when disclosing plaque threshold at level 2, for both dentist RFT-QH evaluation (sensitivity 71.1%, specificity 67.7%, agreement 70.2%) and computer RFT-QH classification (sensitivity 68.4%, specificity 62.9%, agreement 67.1%).

### 3.4 Plaque Progression Between Visits

Figure 6 shows tooth level plaque progressions described by the dentist diagnosis and computer analysis during the clinical study. Gradual increase was found in human visual assessment of dental plaque. Similarly, the average values of computer features increased with the visit number. Taking significant level at  $p < 0.05$ , statistically significant differences were detected between baseline, 24 h, and 48 h visits in the clinical disclosing T-QH scores, dentist RFT-QH scores, and computer RFT-QH scores at tooth and subject levels. For computer features, plaque areas ( $\delta$ ) showed significant differences between any consecutive visits. But there was no significant difference in plaque intensity ( $\Delta P_{ave}$ ) between baseline and 24 h ( $p = 0.076$ ) as well as 24 h and 48 h ( $p = 0.07$ ) at tooth level and between 24 h and 48 h ( $p = 0.15$ ) at subject level.

Because pooling cases (excess fluorescence dye not completely dried) are likely to be scored  $T - QH = 1$  but less clinically relevant, this study took the disclosing score  $T - QH \geq 2$  and red fluorescence index  $RFT - QH \geq 1$  as plaque affected. Figure 7 shows plaque progressions in terms of affected tooth



**Fig. 4** Boxplot describing the distributions of computer features at different red fluorescence plaque levels. Red line in each box represents the median value of the feature in each class. Box boundaries in blue lines show the 25% quartile (lower) and 75% quartile (higher) of the data. Boundaries in black lines describe the extreme values in the tails of the feature distribution to rule out the outliers. We set whisker parameter to 1 in this case. Red cross beyond black lines is the outlier. (a)–(c) Distribution of computer features based on dentist plaque assessment. (d)–(f) Distribution of computer features based on computer plaque classification. (a) and (d) Average plaque intensity  $\Delta P_{ave}$ . (b) and (e) Plaque areas  $\delta$  (%). (c) and (f) Inner product of plaque intensity and plaque areas  $\Delta P_{ave} * \delta$ .

**Table 3** Spearman rank correlation between clinical disclosing plaque scores and red fluorescence RFT-QH index at tooth and subject levels. Correlations without superscript “\*” are significant at 0.01 level (two-tailed).

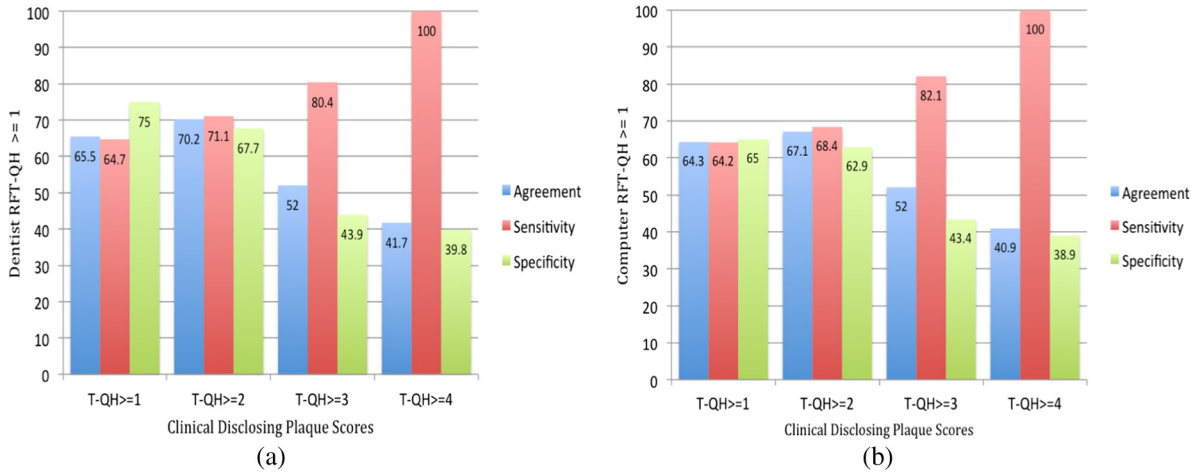
Clinical T-QH	Tooth level		Subject level	
	Dentist RFT-QH	PC RFT-QH	Dentist RFT-QH	PC RFT-QH
BL	0.284	0.251	0.583	0.641
24 h	0.369	0.430	0.424*	0.479*
48 h	0.367	0.342	0.548	0.541
All visits	0.450	0.392	0.618	0.547

number per subject by human visual evaluation and computer classification. It is noted that the affected tooth number per subject increased with time for both disclosing scores and red fluorescence scores. After 48 h with no oral hygiene, all the subjects had six teeth affected with plaque, which was detectable by disclosing solutions in clinical assessment. Red fluorescence plaque was not identified on all the tooth surfaces. But for subjects at 48 h appointment, there were 10 out of 14 participants having at least five teeth affected with red fluorescence plaque by dentist diagnosis. Computer analysis showed lower plaque level, giving eight subjects in this case.

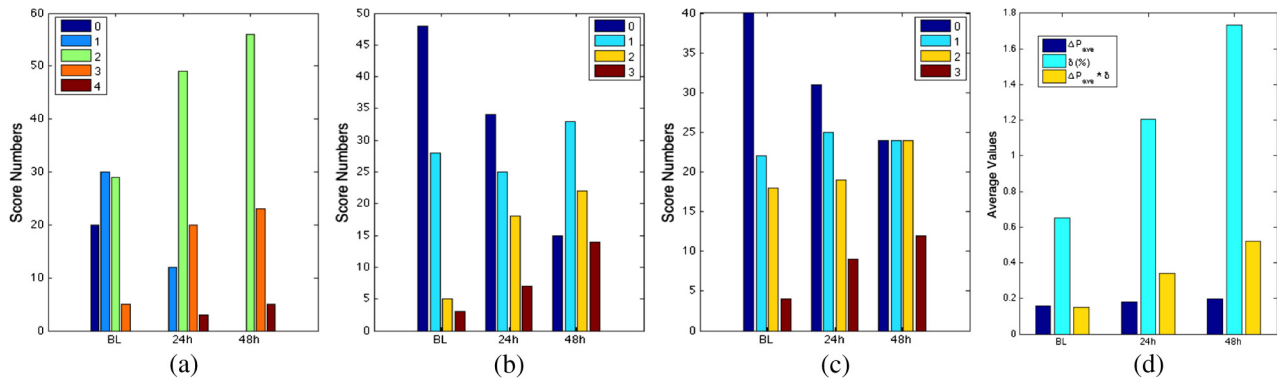
**Table 4** Spearman rank correlation between computer features and human visual assessment at tooth and subject levels. Correlations without superscript “\*” are significant at 0.01 level (two-tailed).

Computer features		Tooth level		Subject level	
		Clinical T-QH	Dentist RFT-QH	Clinical T-QH	Dentist RFT-QH
BL	$\Delta P_{ave}$	0.176*	0.348	0.250*	0.622
	$\delta$ (%)	0.289	0.613	0.738	0.741
	$\Delta P_{ave} * \delta$	0.286	0.613	0.722	0.701
24 h	$\Delta P_{ave}$	0.357	0.473	0.313*	0.682
	$\delta$ (%)	0.437	0.659	0.466*	0.731
	$\Delta P_{ave} * \delta$	0.443	0.656	0.488*	0.724
48 h	$\Delta P_{ave}$	0.329	0.652	0.520*	0.816
	$\delta$ (%)	0.377	0.818	0.597	0.903
	$\Delta P_{ave} * \delta$	0.396	0.824	0.638	0.767
All visits	$\Delta P_{ave}$	0.300	0.491	0.379	0.746
	$\delta$ (%)	0.397	0.698	0.517	0.764
	$\Delta P_{ave} * \delta$	0.395	0.698	0.534	0.772

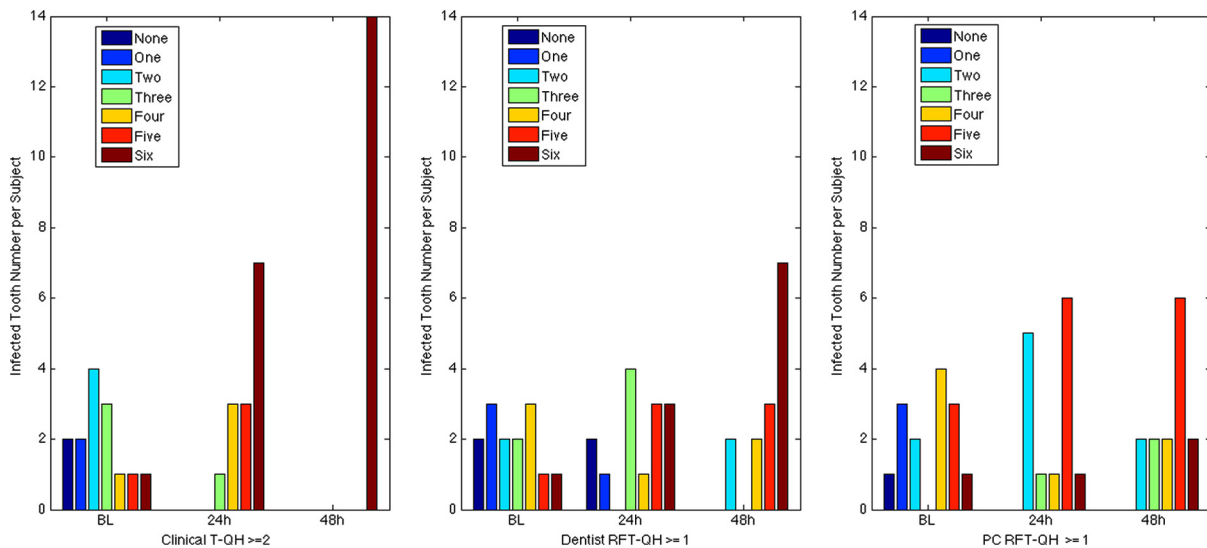




**Fig. 5** Agreement between clinical disclosing plaque scores and red fluorescence plaque scores (RFT-QH). (a) Disclosing plaque scores versus dentist RFT-QH scores. (b) Disclosing plaque scores versus computer RFT-QH scores.



**Fig. 6** Tooth level plaque progressions described by dentist plaque scores, computer plaque scores, and computer features during the clinical study. (a) Case number of disclosing T-QH scores at different levels across the visits. (b) Case number of dentist RFT-QH scores at different levels across the visits. (c) Case number of computer RFT-QH scores at different levels across the visits. (d) Average values of computer red fluorescence features at different visits.



**Fig. 7** Plaque progressions in terms of affected tooth number per subject evaluated by (a) clinical disclosing plaque, (b) dentist red fluorescence plaque, and (c) computer red fluorescence plaque.

### 3.5 Post-hoc Power Analysis

A *post-hoc* power analysis was performed to outline the sample size in the future clinical studies for differentiating plaque level at baseline from that after 48 h plaque accumulation, based on the distributions of computer plaque scores derived from this pilot study. Taking the significant level  $p < 0.05$  (Wilcoxon signed-rank test, two tails) and statistical power  $> 0.9$ , the sample size is minimum 72 at tooth level with effect size of 0.42 and minimum 42 at subject level with effect size of 0.56.

## 4 Discussion

The amount of red fluorescent plaque increased in time when participants refrained from oral hygiene continuously for 48 h after prophylaxis. Statistically significant differences were identified in red fluorescence plaque scores (RFT-QH) and computer features between baseline and 48 h appointment at tooth and subject levels ( $p < 0.05$ ). This indicates that the plaque progression could be monitored using red fluorescence imaging technique.

Taking clinical diagnosis as reference standard, moderate correlations showed between clinical disclosing indices and red fluorescence scores. In this pilot study, ~30% teeth (false-negative cases), with young plaque revealed by disclosing solutions, showed no red fluorescence plaque in the images. This is probably because there was an insufficient amount of porphyrins in these young plaques and thus little signals presented in the red fluorescence images. Hence, red fluorescence plaque generally believed as a property of matured plaque is supported by the outcome of this study. Since young plaque is considered relatively healthy<sup>10</sup> while old plaque leads to the development of tooth decay and gum diseases, red fluorescence imaging could provide an objective measure for evaluating the oral health risks.

On the other hand, there were 30% to 40% teeth (false-positive cases) that showed red fluorescence signals in the images but had no disclosing plaque discovered in the clinical examination. These incidents included several tooth defects, such as caries, calculus, hypomineralization, and discoloration, some of which were reported detectable by the red fluorescence imaging.<sup>4,5,9</sup> Possible confounders were included within the study because this work was interested in the validity of fluorescence signal for use by a lay population, who would not easily be able to discriminate the source of the signals. Considering red fluorescence signal was similar for plaque and other types of tooth defects, separation of signals from sources other than plaque through image processing is challenging and one of the future focuses of research moving forward. In addition, because this study only scaled and cleaned target teeth, a number of false-positive cases resulted from matured plaque remaining at the interproximal areas between adjacent teeth. This deficiency in prophylaxis reflected by the dental plaque exposed at baseline, where only two subjects had no plaque identified in the clinical assessment.

There were great variations on the individual responses of red fluorescence plaque after 48 h accumulation without oral hygiene. It was reported that elderly people were more likely to become heavy plaque former than young people.<sup>23</sup> Hence, it would be worthwhile to categorize individuals according to their responses of red fluorescence plaque. The analysis of the common characteristics of individuals within each group could be helpful for understanding the influential factors of plaque formation.

For computer analysis, affected plaque areas correlated well with human visual assessment on red fluorescence images, whereas weak to moderate correlations were found in average plaque intensity. This indicates that plaque progression is more obvious in terms of plaque area involvement than redness enhancement. Automatic fluorescence scores showed comparable results with the dentist RFT-QH evaluation when associating with clinical disclosing indices. So, it could be feasible to develop an automatic platform for the quantification of fluorescence plaque when the dentist evaluation was not available.

This pilot study is based on the fact that a limited number of participants were included and only one dentist's evaluation was used to create the clinical references. As such, it is not possible to perform interoperator analysis on an extensive dataset. However, while resources were not available, in our opinion, the study's results still elucidate the usefulness and efficiency of the fluorescence imaging for proper oral hygiene assessment. Nevertheless, future studies would be conducted on a large population based on the *post-hoc* power analysis to investigate whether the derived results are on the order of inter-rater outcomes. Additional consideration will be given to potential confounders via separating the fluorescence signals from sources using image-processing methods. Moreover, subsequent experiments will perform a complete prophylaxis procedure including adjacent teeth to understand its influence on specificity for fluorescence plaque detection.

In conclusion, this paper introduced a red fluorescence imaging system allowing full access into the human mouth and demonstrated its usefulness for dental plaque detection and quantification through a pilot study. Given the correlation with clinical evaluation, red fluorescence imaging shows its potential for providing an objective and promising method for proper oral hygiene assessment. The results of the current study present evidence to support the future work for running a comprehensive experiment on a large dataset.

### Disclosures

This study was sponsored by a research grant from Colgate Palmolive Company. Dr. Juliana Gomez, Ms. Debbie Peru, Ms. Soniya Khan, and Prof. Roger Ellwood are employees of Colgate Palmolive Company.

### Acknowledgments

The authors would like to thank Ms. Sylvia L. Santos for performing the clinical examinations and Ms. Elizabeth Gittins for the arrangement of the clinical studies.

### References

1. P. D. Marsh, "Dental plaque as a biofilm and a microbial community-implications for health and disease," *BMC Oral Health* **6**, S14 (2006).
2. M. Cugini, M. Thompson, and P. R. Warren, "Correlations between two plaque indices in assessment of toothbrush effectiveness," *J. Contemp. Dent. Pract.* **7**, 1–9 (2006).
3. R. Heinrich-Weltzien et al., "Quantitative light-induced fluorescence (QLF™): a potential method for the dental practitioner," *Quintessence Int.* **34**(3), 181–188 (2003).
4. Q. G. Chen et al., "Quantitative method to assess caries via fluorescence imaging from the perspective of autofluorescence spectral analysis," *Laser Phys.* **25**, 085601 (2015).
5. L. Karlsson, "Caries detection methods based on changes in optical properties between healthy and carious tissue," *Int. J. Dent.* **2010**, 270729 (2010).

6. I. A. Pretty et al., "Quantification of dental fluorosis using fluorescence imaging," *Caries Res.* **40**, 426–434 (2006).
7. C. K. Hope et al., "Photobleaching of red fluorescence in oral biofilms," *J. Periodontal Res.* **46**, 228–234 (2011).
8. M. H. Van der Veen, "Red autofluorescence of dental plaque bacteria," *Caries Res.* **40**, 542–545 (2006).
9. E. Terrer et al., "A new concept in restorative dentistry: light-induced fluorescence evaluator for diagnosis and treatment. Part 1. Diagnosis and treatment of initial occlusal caries," *J. Contemp. Dent. Pract.* **10**, E086–E94 (2009).
10. M. H. Van der Veen et al., "Dynamics of red fluorescent dental plaque during experimental gingivitis: a cohort study," *J. Dent.* **48**, 71–76 (2016).
11. P. Rechmann et al., "Performance of a light fluorescence device for the detection of microbial plaque and gingival inflammation," *Clin. Oral Invest.* **20**, 151–159 (2016).
12. C. M. Volgenant et al., "Comparison of red autofluorescing plaque and disclosed plaque: a cross-sectional study," *Clin. Oral Invest.* **20**(9), 2551–2558 (2016).
13. M. Oren and S. K. Nayar, "Generalization of Lambert's reflectance model," in *Proc. of the 21st Annual Conf. on Computer Graphics and Interactive Techniques*, pp. 239–246 (1994).
14. C. Seiffert et al., "RUSBoost: a hybrid approach to alleviating class imbalance," *IEEE Trans. Syst. Man Cybern.—Part A: Syst. Humans* **40**(1), 185–197 (2010).
15. J. Friedman, T. Hastie, and R. Tibshirani, "Additive logistic regression: a statistical view of boosting," *Ann. Stat.* **28**(2), 337–407 (2000).
16. J. R. Quinlan, *C4.5: Programs for Machine Learning*, Morgan Kaufmann Publishers, San Francisco, California (1993).
17. Y. Freund and R. E. Schapire, "Experiments with a new boosting algorithm," in *Proc. of 13th Int. Conf. of Machine Learning*, pp. 148–156 (1996).
18. G. Seymour, *Predictive Inference*, Chapman and Hall, New York (1993).
19. T. Fawcett, "An introduction to ROC analysis," *Pattern Recognit. Lett.* **27**(8), 861–874 (2006).
20. J. Cohen, "Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit," *Psychol. Bull.* **70**, 213–220 (1968).
21. A. Lehman, *JMP for Basic Univariate and Multivariate Statistics: A Step-by-Step Guide*, p. 123, SAS Press, Cary, North Carolina (2005).
22. K. P. Suresh and S. Chandrashekar, "Sample size estimation and power analysis for clinical research studies," *J. Hum. Reprod. Sci.* **5**(1), 7–13 (2012).
23. C. Fransson, T. Berglundh, and J. Lindhe, "The effect of age on the development of gingivitis, clinical, microbiological and histological findings," *J. Clin. Periodontol.* **23**(4), 379–385 (1996).

Biographies for the authors are not available.