Diagnosis of oral lichen planus from analysis of saliva samples using terahertz time-domain spectroscopy and chemometrics

Yury V. Kistenev
Alexey V. Borisov
Maria A. Titarenko
Olga D. Baydik
Alexander V. Shapovalov

SPIE.

Diagnosis of oral lichen planus from analysis of saliva samples using terahertz time-domain spectroscopy and chemometrics

Yury V. Kistenev,[a,b,*] Alexey V. Borisov,[a,b] Maria A. Titarenko,[c] Olga D. Baydik,[c] and Alexander V. Shapovalov[a]
[a]Tomsk State University, Laboratory of Biophotonics, Tomsk, Russia
[b]Siberian State Medical University, Department of Physics and Mathematics, Tomsk, Russia
[c]Siberian State Medical University, Department of Stomatology, Tomsk, Russia

Abstract. The ability to diagnose oral lichen planus (OLP) based on saliva analysis using THz time-domain spectroscopy and chemometrics is discussed. The study involved 30 patients (2 male and 28 female) with OLP. This group consisted of two subgroups with the erosive form of OLP ($n = 15$) and with the reticular and papular forms of OLP ($n = 15$). The control group consisted of six healthy volunteers (one male and five females) without inflammation in the mucous membrane in the oral cavity and without periodontitis. Principal component analysis was used to reveal informative features in the experimental data. The one-versus-one multiclass classifier using support vector machine binary classifiers was used. The two-stage classification approach using several absorption spectra scans for an individual saliva sample provided 100% accuracy of differential classification between OLP subgroups and control group. © *2018 Society of Photo-Optical Instrumentation Engineers (SPIE)* [DOI: 10.1117/1.JBO.23.4.045001]

1 Introduction

Oral lichen planus (OLP) is a chronic, mucocutaneous inflammatory disease with cell-mediated immune dysfunction, hereditary, allergic, neuroendocrine, infectious, and metabolic factors.[1–4] There are six types of OLP: reticular, papular, atrophic, plaque, erosive, and bullous forms.[5] OLP lesions appear most often on the buccal mucosa, retromolar area, and gingival and lateral borders of the tongue, almost always in a bilateral disposition.[6] The lesions are consistently more resistant to treatment than dermal lesions.[7]

OLP is clinically associated with possible transformation to oral cancer.[5,8] The World Health Organization describes OLP as a "generalized state associated with a significant increased risk of cancer."[9] Thus, early detection of OLP plays an important role in successful treatment, prognosis assessment, and stabilization of the disease in remission.[10]

Histological examination of tissue remains the gold standard to identify malignant oral lesions according to morphological criteria.[11] But the biopsy analysis is invasive based on subjective interpretation of the pathologist. Furthermore, it has difficulties in selection of the most appropriate site of biopsy when the lesions are large. This necessitates developing more suitable diagnostic methods.[12]

Cytological examination of mucous membrane cells has known advantages in diagnosis of malignant transformation.[13–19] Remmerbach et al.[15–17] showed that exfoliative cytology methods combined with DNA image cytometry achieved 98.2% accuracy for sensitivity and 100% accuracy for specificity.[18] Exfoliative cytology is a simple, noninvasive, inexpensive method of diagnosis, which if necessary can be repeated with minimal discomfort to the patient. This method also enables detecting the disease at an early stage, monitoring the course of disease, and evaluating the possibility of malignant transformation of OLP lesions *in situ*. However, cytological tests are time-consuming and applicable when evident clinical changes appear.

Commercially available chemiluminescence kits provide noninvasive analysis of oral cavity disorders. But significant reduction of fluorescence can appear in frequently occurring conditions, such as mucosal pigmentations, ulcerations, irritations, gingivitis, and a high concentration of hemoglobin in blood. Chemiluminescence methods also can have difficulties in differential diagnosis. For example, the VELscope imaging system has shown usefulness in confirming the presence of oral leukoplakia, erythroplakia, and other oral mucosal disorders, but the device was unable to discriminate high-risk from low-risk lesions.[20] The ViziLite chemiluminescence detection kit detects potentially malignant oral disorders, can be used as a general oral mucosal examination system, and may, in particular, improve the visualization of leukoplakia, but it does not accurately delineate dysplastic lesions.[21]

Photodynamic diagnosis (PDD) is based on the fluorescence generated by administering an exogenous photoactivated substance that accumulates in tumor cells. One of the most promising photosensitizers for oral cancer diagnosis is 5-aminolevulinic acid (ALA), which can induce protoporphyrin IX fluorescence in tissue. PDD in the oral cavity can be implemented simply by rinsing with a 0.4% ALA solution. A high sensitivity but limited specificity of PDD in the detection of oral cancer has been shown in Ref. 22.

Raman spectroscopy (RS) is successful in differentiating normal tissue from nonoral premalignant and malignant tissues.

In-vivo RS has shown efficacy in detecting malignancy-associated changes in the oral cavity. However, the clinical applications of RS are limited by both the difficulty of capturing weak Raman signals from tissue and the relatively slow speed of spectrum acquisitions.[22]

Elastic scattering spectroscopy was shown to be sensitive to nuclear size, chromatin content, and nuclear/cytoplasmic ratio, which are of interest for detecting malignant tissue, but it is rarely used in premalignant and malignant oral tissue studies.[22]

Diffusive reflectance microscopy allows analyzing tissue morphology, such as nuclear size distribution, epithelial thickness, collagen content, and the amount of oxy- and deoxyhemoglobin, all of which can vary during carcinogenesis in the epithelia. Confocal reflectance microscopy can provide detailed images of the tissue structure and cellular morphology of living tissue in real time.[22]

Optical coherence tomography (OCT) produces cross-sectional images of tissue with spatial resolution up to 10 to 20 μm and in depths up to 1 to 2 mm, that is suitable for study of oral mucosal lesions. However, OCT does not provide quantitative information; also, only a small area can be examined at a time.[22]

Saliva content analysis for providing indirect OLP diagnosis seems to be very attractive. Saliva is produced by the three major salivary glands (parotid, submandibular, and sublingual) as well as by numerous minor salivary glands.[23,24] Water is the main component of saliva (99%). Organic substances are represented by more than 400 types of proteins; the most significant ones are glandular in origin (alpha amylase, histamines, cystatins, lactoferrins, lysozymes, mucins), proline-rich proteins, or plasma derivatives (albumin, secretory immunoglobulin A). The inorganic components of saliva are mainly composed of ions, including Na+, Cl−, Ca_2+, Mg_2+, $(HPO_2)^{3+}$, and NH^3.[25] Saliva also contains peptides, mRNA, DNA, microRNA, and substances and other molecules derived from the systemic circulation.[26]

The noninvasive, simple nature of saliva collection and no necessity for special equipment for sampling are attractive for diagnosis and for monitoring disease progression and treatment responses with minimally trained personnel.[27–29] The use of saliva also offers a cost-effective approach for large-scale screening.[30] For example, salivary proteomic and genomic biomarkers have been successfully examined for the detection of oral squamous cell carcinoma, Sjögren's syndrome, pancreatic cancer, and breast cancer.[31–33]

Ogden et al.[34] suggested that a quantitative technique, based on the evaluation of parameters, such as nuclear area (NA) and cytoplasmic area (CA), may increase the sensitivity of exfoliative cytology for early diagnosis of oral cancers. Cowpe et al.[35] found that as a rule, tissues undergoing malignant transformation show a reduction of CA, and then a reduction of NA. These products should be in saliva.

The nature of emergence of a large number of products of cell destruction in the oral cavity of patients with the erosive form of OLP can be associated with the oxidative stress that accompanies chronic inflammation. The oxidative stress leads to accumulation and activation of leukocytes; increased production of reactive oxygen species; increased production of glutathione and thiobarbituric acid reactive substances; and increased production of C-reactive protein, nitric oxide, nitrates, and nitrites in saliva.[36,37] This variation in the molecular content of saliva can be used for noninvasive express diagnosis of OLP.

The diagnostic potential of saliva is attributed to its molecular profile, which can be illuminated by spectroscopy methods. The most informative is chromatography, but it is time- and money-consuming for routine practice. Laser spectroscopy has rather optimal characteristics for medical applications. The terahertz (THz) spectral range is often called the area of "structural fingerprint." THz laser spectroscopy can analyze intra- and intermolecular interactions; the THz region also contains absorption lines associated with rotation and low-frequency vibration modes of biological macromolecules and deformations of hydrogen bonds. Currently, the results of the study of absorption spectra in the THz region of peptides, proteins, and nearly all amino acids have been obtained.[38–41]

This work evaluates the potential of an approach to OLP diagnosis based on express saliva analysis by laser THz spectroscopy and chemometrics.

2 Materials and Methods

2.1 Groups under Study

The research was carried out according to the principles of good clinical practice. Protocol of the research was approved by the Ethics Committee of the Siberian State Medical University (Tomsk, Russia), Ref. Number 4532 on July 07, 2015. All participants were informed about the details of the research and signed the Informed Agreement. The interaction with the patients was limited to the sampling by a saliva probe placed into a disposable container and the scraping of cell material by a sterile dental metal spatula from the mucous membrane of the cheek for cytological analysis.

The study excluded patients with lichenoid reactions, drug reactions, and systemic lupus erythematosus; patients with restorations of amalgam and acute inflammatory diseases of oral cavity (abscesses, pericoronitis, periodontitis); and patients who were receiving drugs that affect immune response (including oral contraceptives, corticosteroids, retinoids, tacrolimus, cyclosporine, and azathioprine) at the time of the study and for 6 months prior, as well as pregnant or nursing women.

The study participants involved 30 patients (2 males and 28 females) with clinically and histologically proven OLP. This group included two subgroups with the erosive form of OLP ($n = 15$) and with the reticular and papular forms of OLP ($n = 15$). In all cases, the diagnosis of OLP was confirmed clinically and by histopathological examination.

The control group consisted of six healthy volunteers (one male and five females) without inflammation in the mucous membrane in the oral cavity and without periodontitis.

2.2 Sampling Protocol

For 12 h before the sampling, the participants ingested no meals or medications; alcohol and tobacco were also excluded. On the morning of the saliva sampling, toothbrushing, ingestion of food, liquids (including water), or medications, smoking, and chewing gum were excluded. An unstimulated saliva probe was taken on an empty stomach between 8 am and 9:30 am. For the free flow of saliva, the participant tilted his or her chin down to the chest and opened his or her mouth; saliva dripped into the lower lip of a cuvette over 5 min. The cuvette was tightly closed with a cap. The samples were analyzed within 1 h after collection.

Additionally, for every participant, we produced swabs by scraping the mucous membrane of the cheeks with a sterile dental metal spatula; the scraping was then placed on a glass slide and dried for 1 h, fixed in 96% ethanol, and washed in water for 10 min. Smears were stained with Romanovsky–Giemsa stain.

Multiple sampling was used to provide statistical analysis.

2.3 Experimental Base

The stained slides were analyzed by the optical microscope Carl Zeiss Axio Scope (Germany). The scrapings were assessed by the degree of destruction of the epithelial cells. The epithelial cells were divided into the following types: the zero type of cells had a normal nucleus and cytoplasm structure; the first type of cells had no more than 50% damage to the cytoplasm and normal structure of the nucleus; the second type of cells had more than 50% damage to the cytoplasm and partial damage to the nucleus; the third type of cells had complete damage to the cytoplasm but not complete damage to the nucleus; and the fourth type of cells had complete damage and disintegration of both the cytoplasm and the nucleus.

For every slide, 100 epithelial cells were selected and the nuclear/cytoplasmic ratio (N/C) was calculated and each stage of differentiation of epithelial cells was defined according to N/C value. When the N/C was 0.50 to 0.49, a cell was considered at the first stage of differentiation; when N/C was 0.40 to 0.49, a cell was considered in the second stage; when N/C was 0.30 to 0.39, a cell was in the third stage; when N/C was 0.20 to 0.29, a cell was in the fourth stage; and when N/C was 0.10 to 0.19, a cell was in the fifth stage of differentiation. In the absence of the nucleus (N/C = 0), a cell was attributed to the sixth stage of differentiation (Fig. 1). Then, the index of cellular differentiation (ICD) was calculated according to the formula:

$$ICD = 1a + 2b + 3c + 4d + 5e + 6f. \quad (1)$$

Here, the numbers 1 to 6 denote the differentiation stages and also are weighting factors of the stages. The latin letters "a," "b," "c, "d," "e," "f" denote the number of cells corresponding to the specific stage of differentiation.
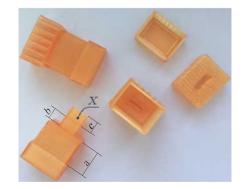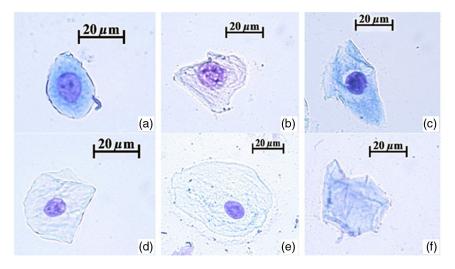
Fig. 2 The cuvette designed for saliva sampling.

Callimeri and Smith[42] found that increase in the N/C was an indicator of the transformation from a benign cell to a malignant one. Cowpe et al.[25] showed that decrease in cytoplasmic diameter, and hence cytoplasmic perimeter and CA, was an early indicator of malignant cell transformation. They found that decrease in the cell diameter occurred in cell dysplasia.
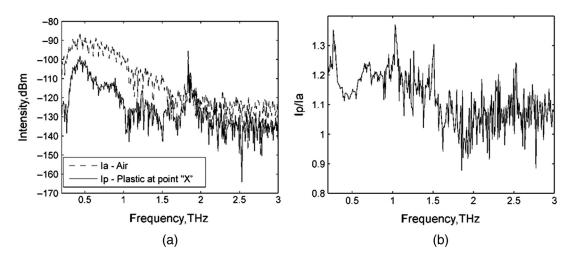
The spectral analysis of saliva samples was produced by the time-domain THz spectrometer T-SPEC (EKSPLA, Estonia). The main characteristics of this spectrometer are a spectral range of 0.3 to 3.5 THz, a dynamic range of up to 90 dB, and spectral resolution of at least 2.3 GHz. To increase the signal-to-clutter ratio, the average of 1024 scans was used.

Special cuvettes were designed for saliva sampling (Fig. 2). The cuvette had a part for measurements ("b" in Fig. 2), which provided a saliva layer thickness of 1 mm, because a thicker layer of saliva almost totally absorbs the THz radiation. The bottom part "c" of the cuvette was designed to produce the reference signal. Tests have shown that the cuvettes are transparent enough in the THz range (Fig. 3). Figure 3(a) shows the intensity of output THz signal when the radiation passed through room air without the cuvette (Ia) and when the radiation passed through the cuvette (Ip) in the point "X" (Fig. 2).

The measurements of output signal at five points on the cuvette surface were performed. This procedure was repeated for 20 cuvette copies.

Fig. 1 Various stages of differentiation of epithelial cells: (a) a cell considered to be at the first stage of differentiation; (b) a cell in the second stage; (c) a cell in the third stage; (d) a cell in the fourth stage; (e) a cell in the fifth stage; and (f) a cell in the sixth stage.

Fig. 3 The intensity of output signals for when the THz radiation passed through room air without the cuvette (Ia) and when THz radiation passed through both (a) room air and the cuvette at the point "X" (Ip) and (b) the ratio of these signals.

2.4 Data Preprocessing and Analysis

The THz absorption spectra of saliva samples were considered as feature vectors of the participant's state. Principal component analysis (PCA) was used to select informative features in the initial feature vectors. The basic idea of PCA is to find a reduced number of new variables that are adequate for recovery of the initial variables, possibly with insignificant errors. The basic idea of PCA consists in projection of correlate variables into a lower number of uncorrelated variables called principal components (PCs).[43]

The supervised support vector machine (SVM) binary classification method with Gaussian radial basis function kernel $K(x, y) = \exp(-\frac{\langle (x-y),(x-y) \rangle}{2\sigma^2})$ or polynomial kernel $K(x, y) = (x^T y + 1)^\rho$ was used for dichotomous classification of the groups under study.[44] Here, x, y are the feature vectors, σ is the optimization parameter, T is the operation of matrix transposition, and ρ is the degree of the polynomial.

There are several approaches to constructing a multiclass classifier using binary classifiers.[45] According to the one-versus-all method, we had to construct N independent binary classifiers, so that every classifier would separate a specific class's feature vectors from all other classes' feature vectors.[46] According to the one-versus-one (also known as all-versus-all) method, we had to construct $N(N-1)$ independent binary classifiers, each of which would separate i'th class feature vectors from j'th class feature vectors.[47] The latter method was shown to provide the better results.[48]

Nonparametric Mann–Whitney U test was also used, and $p < 0.05$ was considered statistically significant.

3 Results

3.1 Cytological Analysis

The cytological analysis was carried out as follows. The number of participants in the control group, the subgroup of patients with the erosive form of OLP, and the subgroup of patients with reticular and papular forms of OLP were denoted as N_1, N_2, N_3, respectively. Then, the indexes $j_1 = 1, 2, \ldots, N_1$, $j_2 = 1, 2, \ldots, N_2$, and $j_3 = 1, 2, \ldots, N_3$ denoted participants in corresponding group.

A swab was taken from every participant, then the stained slide was prepared, from which the random statistical sampling of 100 cells was analyzed. The cells in a sample were marked by $\alpha = 0, 1, \ldots, 4$ depending on type. Then, the number of cells of type α in a swab from patient j_1 from the first group was denoted as $n_{j_1 \alpha}^{(1)}$, the same for the second group was $n_{j_2 \alpha}^{(2)}$, etc.

The results of cytological analysis are presented in Table 2 in terms of the total number of cells of a specific type in the specific group. For example, $A_{1\alpha} = \sum_{j_1=1}^{N_1} n_{j_1 \alpha}^{(1)}$ is the total number of cells of type α in the first group. Its relative value (in %) is $\frac{A_{1\alpha}}{N_1 \times 100} \times 100\%$. In Table 2, p_{ij} is the significant level of difference between average characteristics of i'th and j'th groups according to χ^2 Pearson criterion.

According to cytological studies, the number of epithelial cells of the 0 type of destruction in the control group was 61.2%, whereas for patients with erosive, reticular, and papular forms of OLP, this parameter was decreased significantly to 34.7% and 54.0%, respectively ($p < 0.05$) (Table 1). For the erosive form of OLP, zero type and the fourth type of epithelial cells were observed almost equally (34.7% and 31.5%, respectively). Despite the fact that the zero type and the first type cells predominated for the group with reticular and papular forms of OLP, content of these cells was significantly lower in comparison with the control ($p < 0.05$).
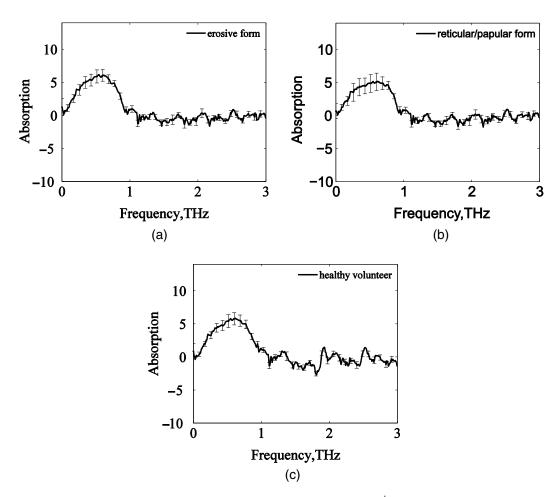
The mean value of the ICD was 448.4 ± 5.7 in the control group and was 397.4 ± 7.7 and 438.6 ± 8.3 in the group with the erosive form of OLP and in the group with the reticular and papular forms of OLP, respectively. Thus, the mean value of the ICD for the reticular and papular forms group showed no statistically significant distinction compared to the control group.

3.2 Spectral Analysis of Saliva Samples

The shapes of the averaged absorption spectra of the saliva samples for patients with erosive OLP, reticular and papular OLP, and healthy volunteers processed by the median filter are shown in Fig. 4. The log-ratio of the output intensities $\log \frac{I_c}{I_b}$ was used to exclude the influence of the cuvette. Here, I_b and I_c are the output intensities of THz radiation that passed through parts "b" and "c" of the cuvette, respectively.

Table 1 Degree of destruction of buccal epithelial cells in the groups under study.

Types of epithelial cells	The degree of destruction of epithelial cells, absolute value and relative value (in %)			p
	Control ($n = 6$)	Erosive form ($n = 15$)	Reticular and papular forms ($n = 15$)	
Zero type	367 (61.2%)	520 (34.7%)	810 (54.0%)	$p_{12} < 0.001$, $p_{13} = 0.04$, $p_{23} < 0.001$
First type	111 (18.5%)	158 (10.5%)	263 (17.5%)	$p_{12} < 0.001$, $p_{13} = 0.588$, $p_{23} < 0.001$
Second type	47 (7.8%)	42 (2.8%)	113 (7.5%)	$p_{12} < 0.001$, $p_{13} = 0.147$, $p_{23} < 0.001$
Third type	39 (6.5%)	323 (21.5%)	180 (12%)	$p_{12} < 0.001$, $p_{13} < 0.001$, $p_{23} < 0.001$
Fourth type	36 (6.0)	473 (31.5%)	135 (9.0%)	$p_{12} < 0.001$, $p_{13} < 0.001$, $p_{23} < 0.001$

(a)

(b)

(c)

Fig. 4 The dependence of the absorption of THz radiation in terms of $\log \frac{I_c}{I_b}$ on the frequency for (c) the control group, (a) patients with erosive OLP, and (b) reticular and papular OLP. The confidence intervals are shown. Here, I_b and I_c are the output intensities of THz radiation that passed through parts "b" and "c" of the cuvette, respectively (see Fig. 2).

The quality of spatial separation of the groups under study in the PCs' space strongly depends on the quantity of the PCs used. We took into account the first eight PCs. The example of the projection of the objects under study on the plane of the second and third PCs is presented in Fig. 5. Evidently, the entire healthy control group is spatially separated from the groups of OLP patients.

To construct a multiclass classifier, we used a set of one-versus-one binary SVM classifiers. The selection of the training subset from all groups under study was repeated randomly many times. The remainder of the initial data was used to test the classifiers. The ratio of separation of initial data on training and testing subsets varied from 0.5/0.5 to 0.85/0.15. A value of the training subset close to 25% from initial data was shown to be optimal.

At first, we constructed the following binary SVM classifiers: patients with the erosive form of OLP versus patients
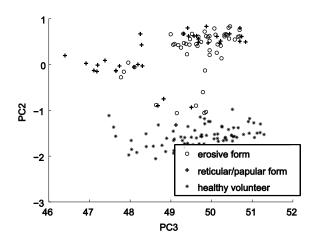
Fig. 5 The projection of the objects under study on the plane of the second and third PCs.

with reticular and papular forms of OLP; patients with the erosive form of OLP versus healthy volunteers; and patients with reticular and papular forms of OLP versus healthy volunteers. The results of testing of the classifiers are presented in Table 2 in the terms of sensitivity and specificity:

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}; \text{Specificity} = \frac{\text{TN}}{\text{FP} + \text{FN}}. \quad (2)$$

Here, TP is the quantity of true-positive classification results and FN is the quantity of false-negative classification results. The results of dichotomous SVM classification with Gaussian radial basis function kernel of the saliva absorption spectra samples for the groups under study are presented in Table 2.

The results of differential (multiclass) diagnosis based on saliva absorption spectra samples analysis and three one-versus-one classifiers from Table 2 are presented in Table 3. The evaluations were carried out using a merged testing set that included patients with erosive, reticular, and papular forms of OLP and healthy volunteers, as shown in Table 3. The feature vector of a representative from the testing set was analyzed by every classifier from Table 2. The differential diagnosis rule was

Table 2 SVM classification of the testing set of saliva absorption spectra samples for the groups under study.

Dichotomous classification	Kernel parameters	Sensitivity		Specificity	
		Mean	Dispersion	Mean	Dispersion
Patients with the erosive form of OLP versus patients with reticular and papular forms of OLP	1.5	0.80	0.12	0.70	0.16
Patients with the erosive form of OLP versus healthy volunteers	2.6	0.95	0.11	0.97	0.10
Patients with reticular and papular forms of OLP versus healthy volunteers	2.4	0.96	0.11	0.97	0.10

Table 3 Differential diagnosis based on the set of SVM classifiers from Table 2.

Group	Quantity of the samples in the testing set	Diagnosis	
		Correct, using single scan for specific sample (%)	Correct, using single sample (four scans for specific sample), %
Patients with the erosive form of OLP	13	88.2	100.0
Patients with the reticular and papular forms of OLP	9	80.0	100.0
Healthy volunteers	19	94.7	100.0

based on the result that was selected in more than 50% of outcomes. The absorption spectrum of every sample was measured four times. Thus, differential classification, based on using a single scan for a specific sample or a single sample (using four scans for a specific sample), was carried out.

In the case of using a single sample, the diagnosis considered to be determined if it coincided for more than two spectra scans for the specific sample under study. This approach in fact represents two-stage classification. An analog of this approach is classification by the "majority vote," when a solution is based on using several classifiers, because using multiple feature vectors for classification of the same object supposes using multiple classifiers. The classification by the majority vote allows increasing the robustness of the decision rules created because of using multiple descriptors for the same object.[49,50]

The results of differential classification based on using saliva samples' absorption spectra scans were shown to be high. The two-stage classification approach based on using several scans for a specific sample provided 100% accuracy in differential classification of the collected data.

4 Conclusion

Cytological cell analysis showed a significant decrease in the ICD in the erosive OLP group when compared to the healthy volunteers and the group with reticular and papular forms of OLP. The decrease is a reflection of the decrease in the N/C of epithelial cells, which is one of the signs of dysplasia. The cytological status of buccal epithelial cells in all patients with OLP indicated disruption in the process of differentiation of the epithelium of the mucous membrane. For patients with erosive OLP, this was characterized by an increased content of epithelial cells with a high degree of destruction.

The differential diagnostic algorithm created based on saliva absorption THz spectra sample analysis and one-versus-one classifiers provided more than 80% accuracy on the test set used. The two-stage classification algorithm described above, which should provide more robust results, in our case, demonstrated 100% accuracy of classification. This will not always be the case; it indicates high homogeneity of the initial data and small overlap of the groups under study in the PCs' space. Owing to this, the two-stage classification algorithm provided excellent filtering of a smaller number of the initial feature

vectors of an object in one group that coincided with feature vectors from the other group.

Histological or cytological cell analysis is rather informative, but simultaneously it is more time-consuming than laser spectroscopy. Automatization of the stained slide preparation is extremely complex and may not be possible in the near future. THz equipment is not very expensive in comparison with routine medical imaging systems. New techiques for creating THz radiation sources and detectors are being designed, for example, silicon CMOS techniques are beginning to extend to the THz domain. Chemical sample analysis using spectral or other approaches will be a basis for evaluating molecular biomarkers in the future.

Disclosures

The authors have no relevant financial interests in the manuscript and no other potential conflicts of interest to disclose.

Acknowledgments

References

1. C. Scully, D. Eisen, and M. Carrozzo, "Management of oral lichen planus," *Am. J. Clin. Dermatol.* **1**(5), 287–306 (2000).
2. J. B. Epstein et al., "Oral lichen planus: progress in understanding its malignant potential and the implications for clinical management," *Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endodontol.* **96**(1), 32–37 (2003).
3. S. B. Ismail, S. K. Kumar, and R. B. Zain, "Oral lichen planus and lichenoid reactions: etiopathogenesis, diagnosis, management and malignant transformation," *J. Oral Sci.* **49**(2), 89–106 (2007).
4. M. S. Alrashdan, N. Cirillo, and M. McCullough, "Oral lichen planus: a literature review and update," *Arch. Dermatol. Res.* **308**(8), 539–551 (2016).
5. J. O. Andreasen, "Oral lichen planus: I. A clinical evaluation of 115 cases," *Oral Surg. Oral Med. Oral Pathol.* **25**(1), 31–42 (1968).
6. B. E. McCartan and C. M. Healy, "The reported prevalence of oral lichen planus: a review and critique," *J. Oral Pathol. Med.* **37**, 447–453 (2008).
7. R. Jayanna et al., "Clinical features, diagnosis and management of oral lichen planus in children," *J. Indian Acad. Oral Med. Radiol.* **25**(3), 209–213 (2013).
8. D. Miricescu et al., "The antioxidant potential of saliva: clinical significance in oral diseases," *Ther. Pharmacol. Clin. Toxicol.* **15**(2), 139–143 (2011).
9. World Health Organization, "Tobacco or health: a global overview," Tobacco or Health: A Global Status Report, 924156184X, pp. 5–65, WHO, Geneva (1997).
10. J. Liu and Y. Duan, "Saliva: a potential media for disease diagnostics and monitoring," *Oral Oncol.* **48**(7), 569–577 (2012).
11. R. Mehrotra et al., "Application of cytology and molecular biology in diagnosing premalignant or malignant oral lesions," *Mol. Cancer* **5**(1), 11 (2006).
12. J. B. Epstein, L. Zhang, and M. Rosin, "Advances in the diagnosis of oral premalignant and malignant lesions," *J. Can. Dent. Assoc.* **68**(10), 617–621 (2002).
13. G. R. Ogden, J. G. Cowpe, and M. W. Green, "Detection of field change in oral cancer using oral exfoliative cytologic study," *Cancer* **68**(7), 1611–1615 (1991).
14. D. Maraki et al., "Very early cytological and DNA cytometric diagnosis of in situ carcinoma in an immunosuppressed liver transplant recipient," *J. Oral Pathol. Med.* **35**(1), 58–60 (2006).
15. T. W. Remmerbach et al., "Cytologic and DNA cytometric early diagnosis of oral cancer," *Anal. Cell. Pathol.* **22**(4), 211–221 (2001).
16. T. W. Remmerbach et al., "Earliest detection of oral cancer using non-invasive brush biopsy including DNA-image-cytometry: report on four cases," *Anal. Cell. Pathol.* **25**(4), 159–166 (2003).
17. T. W. Remmerbach et al., "Diagnostic value of nucleolar organizer regions (AgNORs) in brush biopsies of suspicious lesions of the oral cavity," *Anal. Cell. Pathol.* **25**(3), 139–146 (2003).
18. D. Maraki, J. Becker, and A. Boecking, "Cytologic and DNA cytometric very early diagnosis of oral cancer," *J. Oral Pathol. Med.* **33**(7), 398–404 (2004).
19. J. J. Sciubba et al., "Improving detection of precancerous and cancerous oral lesions: computer-assisted analysis of the oral brush biopsy," *J. Am. Dent. Assoc.* **130**(10), 1445–1457 (1999).
20. K. H. Awan, P. R. Morgan, and S. Warnakulasuriya, "Evaluation of an autofluorescence based imaging system (VELscope™) in the detection of oral potentially malignant disorders and benign keratosis," *Oral Oncol.* **47**, 274–277 (2011).
21. K. H. Awan, P. R. Morgan, and S. Warnakulasuriya, "Utility of chemiluminescence (ViziLite™) in the detection of oral potentially malignant disorders and benign keratosis," *J. Oral Pathol. Med.* **40**, 541–544 (2011).
22. D. Liu et al., "Non-invasive techniques for detection and diagnosis of oral potentially malignant disorders," *Tohoku J. Exp. Med.* **238**, 165–177 (2016).
23. R. J. Turner and H. Sugiya, "Understanding salivary fluid and protein secretion," *Oral Dis.* **8**(1), 3–11 (2002).
24. S. P. Humphrey and R. T. Williamson, "A review of saliva: normal composition, flow, and function," *J. Prosthet. Dent.* **85**(2), 162–169 (2001).
25. J. G. Cowpe, R. B. Longmore, and M. W. Green, "Quantitative exfoliative cytology of abnormal oral mucosal smears," *J. R. Soc. Med.* **81**(9), 509–513 (1988).
26. C. S. Miller et al., "Salivary biomarkers of existing periodontal disease: a cross-sectional study," *J. Am. Dent. Assoc.* **137**(3), 322–329 (2006).
27. W. Liu et al., "IFN-gamma and IL-4 in saliva of patients with oral lichen planus: a study in an ethnic Chinese population," *Inflammation* **32**(3), 176–181 (2009).
28. E. Kaufman and I. B. Lamster, "The diagnostic applications of saliva—a review," *Crit. Rev. Oral Biol. Med.* **13**(2), 197–212 (2002).
29. L. R. Bigler et al., "The potential use of saliva to detect recurrence of disease in women with breast carcinoma," *J. Oral Pathol. Med.* **31**(7), 421–431 (2002).
30. D. T. Wong, "Salivary diagnostics powered by nanotechnologies, proteomics and genomics," *J. Am. Dent. Assoc.* **137**(3), 313–321 (2006).
31. S. Hu et al., "Salivary proteomics for oral cancer biomarker discovery," *Clin. Cancer Res.* **14**(19), 6246–6252 (2008).
32. Y. Li et al., "Salivary transcriptome diagnostics for oral cancer detection," *Clin. Cancer Res.* **10**(24), 8442–8450 (2004).
33. F. Wei et al., "Electrochemical sensor for multiplex biomarkers detection," *Clin. Cancer Res.* **15**(13), 4446–4452 (2009).
34. G. R. Ogden, J. G. Cowpe, and A. J. Wight, "Oral exfoliative cytology: review of methods of assessment," *J. Oral Pathol. Med.* **26**(5), 201–205 (1997).
35. J. G. Cowpe, M. W. Green, and G. R. Ogden, "Quantitative cytology of oral smears. A comparison of two methods of measurement," *Anal. Quant. Cytol. Histol.* **13**(1), 11–15 (1991).
36. D. Darczuk et al., "Salivary oxidative status in patients with oral lichen planus," *J. Physiol. Pharmacol.* **67**(6), 885–894 (2016).
37. A. Tvarijonaviciute et al., "Evaluation of salivary nitric oxide levels and C-reactive protein in patients with oral lichen planus and burning mouth syndrome," *J. Oral Pathol. Med.* **46**(5), 387–392 (2017).
38. P. Bakopoulos et al., "A tunable continuous wave (CW) and short-pulse optical source for THz brain imaging applications," *Meas. Sci. Technol.* **20**, 104001 (2009).
39. E. P. J. Parrott, Y. Sun, and E. Pickwell-MacPherson, "Terahertz spectroscopy: its future role in medical diagnoses," *J. Mol. Struct.* **1006**, 66–76 (2011).
40. A. G. Markelza, A. Roitberg, and E. J. Heilweil, "Pulsed terahertz spectroscopy of DNA, bovine serum albumin and collagen between 0.1 and 2.0 THz," *Chem. Phys. Lett.* **320**, 42–48 (2000).
41. X. Yang et al., "Biomedical applications of terahertz spectroscopy and imaging," *Trends Biotechnol.* **34**, 810–824 (2016).

42. G. E. Callimeri and C. J. Smith, "Exfoliative cytology of early lesions of experimental oral cancer in the hamster," *Arch. Oral Biol.* **10**, 465–470 (1965).

43. L. Pomerantsev and O. Y. Rodionova, "Concept and role of extreme objects in PCA/SIMCA," *J. Chemom.* **28**(5), 429–438 (2014).

44. G. R. G. Lanckriet et al., "Learning the kernel matrix with semidefinite programming," *J. Mach. Learn. Res.* **5**, 27–72 (2004).

45. M. Aly, "Survey on multiclass classification methods," Technical Report, pp. 1–9, California Institute of Technology (2005).

46. X. Zhao, S. Guan, and K. L. Man, "An output grouping based approach to multiclass classification using support vector machines," *Lect. Notes Electr. Eng.* **393**, 389–395 (2016).

47. J. Milgram, M. Cheriet, and R. Sabourin, "'One against one' or 'one against all': which one is better for handwriting recognition with SVMs?" in *10th Int. Workshop on Frontiers in Handwriting Recognition* (2006).

48. L. Yang, B. Jian-Wu, and F. Zhi-Ping, "A method for multi-class sentiment classification based on an improved one-vs-one (OVO) strategy and the support vector machine (SVM) algorithm," *Inf. Sci.* **394–395**, 38–52 (2017).

49. G. James, "Majority vote classifiers: theory and applications," Doctoral Dissertation, p. 123, Stanford University (1998).

50. V. Giorgio and F. Masulli, "Ensembles of learning machines," *Lect. Notes Comput. Sci.* **2486**, 3–20 (2002).

Yury V. Kistenev is the author of more than 130 journal papers, including patents and conference proceedings. His current research interests include application of laser photoacoustic spectroscopy in medicine and biology.

Alexey V. Borisov published more than 50 articles in the fields of nonlinear mathematical physics, living systems, plasma physics, numerical modeling, and biophotonics. His areas of interest and experience are Bose–Einstein condensation, quantum computers, computer vision, medical physics, and biological optics.

Maria A. Titarenko is a postgraduate student. She specializes in noninvasive methods of diagnostics of precancer diseases of oral cavity.

Olga D. Baydik is an expert in the morphological methods of diagnostics and treatment of inflammatory and precancer diseases in maxilla-facial region, including paranasal sinuses, oral cavity. She has published more than 50 papers in refereed journals.

Alexander V. Shapovalov specializes in the mathematical methods of theoretical physics, including integrability problems and symmetry analysis in mathematical physics, soliton theory and its applications, and semiclassical approach in nonlinear mathematical physics. He has published more than 150 papers in refereed journals. His current research is related to applications of mathematical methods in biophysics and medical data analysis.