Nogueira MS, et al., J Pharm Pharmacol Res 2024
DOI:10.26502/fjppr.0101

osteoporosis and impaired fertility, including male impotence [1], [2], [3]. Tobacco smoke contains more than 7000 chemicals that damage various organs and tissues [2], [8]. Chemical dependency related to smoking is a habit that is harmful to oral health and can also cause several diseases such as tooth decay, periodontal disease, halitosis and mouth cancer [9]. Saliva is the first biological fluid to come in contact with tobacco smoke and is responsible for maintaining a healthy balance in the mouth, thus leading to the risk of tooth decay[9], [10]

Biomarkers are immunoreactive compounds found in body fluids, such as saliva, that can indicate normal or abnormal processes in the body. Saliva has been used as a diagnostic tool for systemic and chronic diseases such as neoplasia. For example, the proteins that make up saliva can indicate normal function or the risk of a disease occurring due to changes in the common patterns of proteins [11].

Squamous cell carcinoma (SCC) represents more than 90% of oral cancers [12], [13], [14], [15] and includes extrinsic factors such as smoking status and/or alcohol consumption [2], [8]. Tobacco smoking is the most common cause of SCC and corresponds to approximately 90% of cases. Tobacco carcinogens include nitrosamine and benzene [2], [8]. Currently, cancer is diagnosed by histopathological analysis following biopsies, which detects cancer based on morphological cellular changes. However, this technique has disadvantages such as environmental contamination by the agents used for sample preparation, the time-consuming analysis process and experience of the pathologist. To suppress these disadvantages, optical biopsy techniques [5], [16], [17], [18] have been increasingly studied for the detection of tissue biochemical changes. One of these techniques is Fourier-transform infrared (FT-IR) spectroscopy, which has the potential to be used for disease detection, diagnosis and monitoring by analyzing the chemical composition of cells and tissues.

FT-IR spectroscopy consists of illuminating samples with polychromatic light and measure the amount of infrared light absorbed by these samples. Fourier Transform is used to convert the collected raw data into absorption spectra [19]. The advantages of FT-IR spectroscopy consist of easy instrumentation, minimal sample preparation, small sample volumes, and real-time *in vitro* diagnosis [18], [20]. As an optical vibrational spectroscopy method, it has the ability to non-invasively characterize biomolecules including the observation of lipid, protein, nucleic acid and carbohydrate levels. The characterization of the biomolecules is typically based on the region of FT-IR spectra [21] and can

LV FRPPRQO\XVHG WR GLuHUHQWLDWH KHDOWKDQG SDWKRORJLFDO tissues [20].

FTIR spectroscopy has been researched as a potential alternative to laboratory tests due to high sensitivity and specificity, rapid analysis, and low cost [20]. Also, patient uptake can be facilitated by the non-invasiveness and fast evaluation provided by FT-IR as a throughput screening tests available commercially, patient prognosis for diseases to be detected can be considerably improved. Also, patient uptake can be facilitated by the non-invasiveness and fast evaluation provided by FT-IR as a point-of-care diagnostics, since FT-IR is currently a mature and widely available technology with higher chances to be translated to clinic. Clinical translation is only possible after collecting data from a considerable number of patients and identifying signal features corresponding to early stages of the disease to be detected. In this study, this disease is SCC, which is predominantly caused by smoking. Therefore, with REMHFWLYH RI WKLV VWXG\ WR HYDOXDWH WKH FODVVLILFDWLRQ SHUIRUPDQFH RI PDFKLQHOHDUQLQJ VPSOH DQG FODVVLILFDWLRQ VDQDOLVH RI FRQVLGHUHG (non-smoker), smoker, and occasional smoker groups using FT-IR spectroscopy and to determine their intrinsic molecular FKDQJHVDVZHOODVWHSHUIRUPDQFHRIVDPSOHGLuHUHQW E\XVLQJVHYHUDOWHVWVRIFODVVLILHUV:LWSHUVSHFWLYHWRWH future aim to compare FT-IR spectra of saliva of smoking, cancer patients and health patients to potentially identify features of early-stage SCC.

0DWHULDOVDQG0HWKRGV

The study was approved by the Research Ethics Committee of the University of Taubaté – UNITAU, São Paulo, Brazil, under protocol number 19436919.7.0000.5501. It was conducted in accordance with Helsinki Declaration [22], informed consent was obtained from all subjects prior to their participation in the study. Patients were considered eligible IRUHQUROPHQWDQWKHUHVHDUFKLIWKH1XOOOHGWKHIROORZ inclusion criteria: all volunteers are over 18 years-old and no gender distinction, smoking and occasionally smoker was considered before accepted, saliva samples were collected. Samples were collected at the Department of Dentistry of the University of Taubaté and in a private clinic in Joinville. The 28 volunteers were divided into control non-smoker (n=11), smoker (n=9), and occasional smoker (OS) (n=8) groups. In the control group, saliva samples from 11 volunteers who had never tried any form of tobacco were collected and analyzed. In the occasional smoker group, 8 samples of patients who smoke sporadically and/or socially were collected and analyzed.

Saliva samples were collected by spit/expectoration.

Nogueira MS, et al., J Pharm Pharmacol Res 2024
DOI:10.26502/fjppr.0101

Participants were instructed to rinse water for one minute and remain without swallowing for a few minutes. At the end of the given time, participants spit all saliva into a sterile universal collector (Figure 1A). All samples were vortexed to be homogenized. After performing this process, samples were transferred from the sterile universal biofluid collector to the Eppendorf microtube (Figure 1B) using a pipette calibrated at 1000µm. The samples were stored in a freezer at -20ºC. The samples were transferred in a Styrofoam with ice to the Institute for Energy and Nuclear Research of University of Sao Paulo at the Center for Laser and Applications (CLA) for analysis. Thermo Scientific Nicolet 6700 ATR FT-IR Spectrometer (Figure 1C) was used to collect spectral measurements presented in this study. The spectrometer contains a diamond crystal for the acquisition of spectral data. In the equipment, 1 µl of saliva sample were placed on the crystal (without additives) by using a calibrated pipette. Once samples were placed on the crystal, these samples were allowed to completely dry for an average time of 5 minutes. Drying time ranged from 2 to 7 minutes. The analysis was performed with 32 scanning scans to obtain the average of the spectra with a resolution of 4cm-¹. Spectra were obtained in triplicate.

After each analysis, the sample was removed from the crystal using absorbent paper and cleaned with 92.8% alcohol to totally remove the sample from the crystal and avoid contamination samples to be measured next. 32 background scans of the equipment were performed for each sample analyzed to remove possible instrumental and environmental interferences. The samples were analyzed by the FT-IR spectral fingerprint region between 900-1800 cm^{-1}. In the control group, 11 samples were analyzed and 33 spectra were obtained. In the sporadic smoker group, 8 samples were analyzed and 24 spectra were obtained. In the smoker group, 9 samples were analyzed and 27 spectra were obtained. A total of 84 spectra was analyzed. The spectral pre-processing and analysis were performed in the Origin Pro8.5 and in the Orange software, respectively. The article from Movasaghi et. al. [23] was used to identify the vibrational modes with found in saliva samples and associate them with biochemical compounds. During the pre-processing step, the raw FT-IR spectra was smoothed by using a Savitsky-Golay filter (2nd polynomial order in 11 points of the frame window), baseline corrected by using a Rubber band algorithm with positive peak direction, and vector normalized.

The evaluation of the saliva classification model was performed by using the leave-one-out validation. Classification performance metrics were obtained for the following classifiers k-nearest neighbors (kNN), decision tree, Support vector machine (SVM) with a radial basis function (RBF) kernel, Stochastic Gradient Descent (SGD), Random Forest, Neural Network, Naive Bayes, Logistic Regression, Gradient

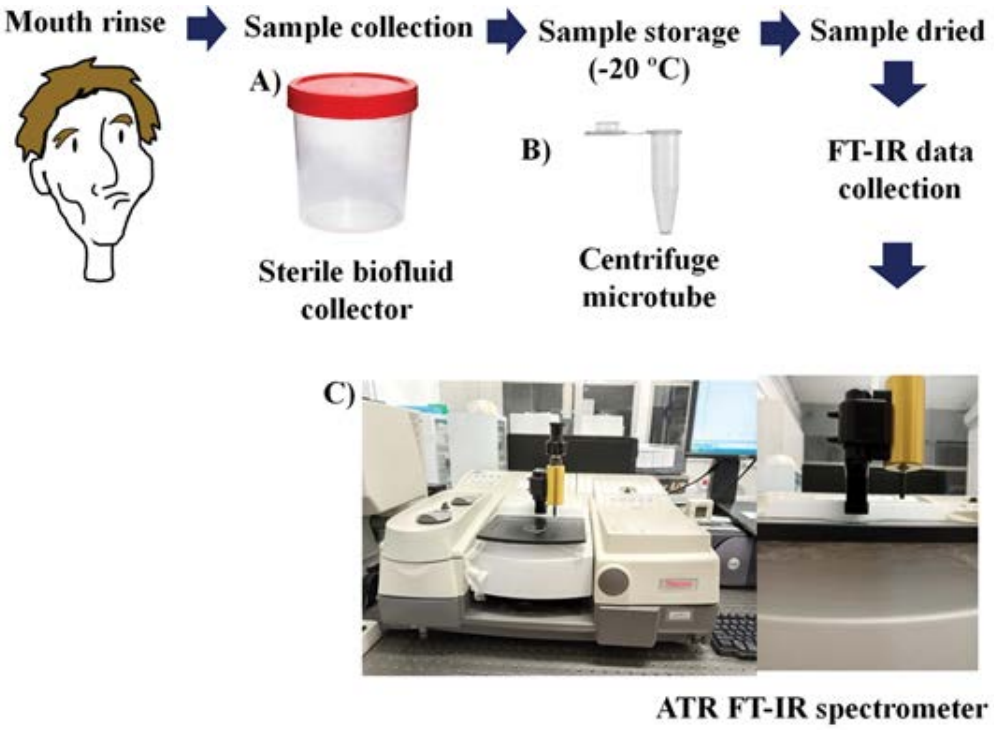


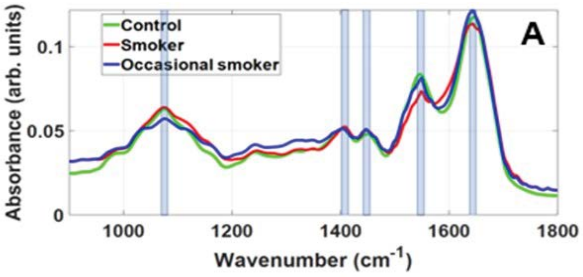
Figure 1: A) Sterile biofluid collector, B) Centrifuge microtube used to store saliva samples in freezer, and C) Thermo Scientific Nicolet 6700 ATR FT-IR spectrometer used for spectral measurements collected in this study.

Nogueira MS, et al., J Pharm Pharmacol Res 2024
DOI:10.26502/fjppr.0101

Boosting and AdaBoost. The classification performance metrics for each classification model were obtained by including all groups (non-smokers, smokers, and occasional smokers) and each pair of groups separating between 1) the control and smokers, 2) smokers and occasional smokers, and 3) control and occasional smokers).

Results

Figure 2 shows that the spectral mean (solid line) and standard deviation (shaded area) of each group overlap considerably. Values for the sporadic smoker group tend to be lower in the region between 1000-1200 cm^{-1} and have a larger standard deviation in this region and that between 1300-1500 cm^{-1}. In the mean spectra (figure 2), some FTIR peaks can be evidenced and assigned to vibrational modes according to Table 1. The peak occurring in 1076 cm^{-1} corresponds to the skeletal cis conformation of DNA and symmetric phosphate [PO_2] stretching. The peak centered in 1403 cm^{-1} is associated with the symmetric CH_3 bending modes of protein methyl groups and $\delta_s CH_3$ of collagen. The peak at 1451 cm^{-1} corresponds to asymmetric CH_3 bending modes of protein methyl groups. The peak at 1547 cm^{-1} corresponds to the amide II group of peptides and proteins, and the peak at 1646 cm^{-1} which corresponds to amide I, C_5 methylated cytosine C==O, stretching C==C uracil, and NH_2 guanine.



Bands (cm^{-1})	Vibrational modes	Structural (Biomolecular) components
1076	Symmetric phosphate [PO_2] stretching	DNA
1403	CH_3 bending of methyl groups (proteins) and $\delta_s CH_3$ (collagen)	Proteins and collagen
1451	Asymmetric CH_3 bending modes of methyl groups	Proteins
1547	Amide II	Peptides and proteins (Amide II)
1646	Amide I, C==O (C_5 methylated cytosine), stretching C==C (uracil), NH_2 guanine	Peptides and proteins (Amide I), methylated cytosine C_5, uracil, and guanine

B

Figure 2: A) Mean and standard deviation of FT-IR spectra collected in this study. The control group is represented by the curve and shaded area in blue color, the sporadic smoker group is represented in red color, and the smoker group is represented in the green color. **B)** Assignment of vibrational modes and structural components to FT-IR peaks evidenced in saliva spectra of the control group.

In table 2, tables A and B indicate the results of the classifier and shows that the most accurate model was the Neural Network. Classification performance metrics of the Neural Network classifier included 0.857 of Area Under the receiver operating characteristic Curve (AUC), classification accuracy of 0.713, F1 score of 0.712, precision of 0.715, and recall of 0.713 (Table A). The confusion matrix (Table B) shows that correctly classified instances were 72.7% for the control group, 65.5% for occasional smokers and 75% for smokers. Tables C, D, E, F, G and H, show that SGD was the most accurate classification model for the group pairs 1) control and smokers, and 2) control and occasional smokers, whereas Neural Network was the most accurate for the group pair smokers and occasional smokers. When classifying samples of the control and smoker groups, 86% of specificity and 84.4% of sensitivity were achieved (F). The classification of control and occasional smoker groups resulted in 76.7% specificity and 73.9% sensitivity (G). Finally, 71.4% of occasional smoker measurements and 78.8% of smoker measurements were correctly classified by the Neural Network model (H).

Table 1: Classifier parameters used to build saliva classification models.

Adaboost	
Base estimator	Tree
Number of estimators	50
Learning rate	1
Classification algorithm	SAMME.R
Regression loss function	Linear
SVM	
Cost	1
Regression loss epsilon 0.1	
RBF Kernel	
Tolerance	0.001
Interation limit	100
Logistic Regression	
Regularization type	Ridge (L2)
Stranght C=1	
SGD	
Classification	Hinge
Regression	Squate Loss
Regularization	Ridge (L2)
Strength (alpha)	0.00001
Learning rate	constant (=0.01)
Number of iterations	1000
Tolerance	0.001
Gradient Boosting	
Number of Trees:	100

Nogueira MS, et al., J Pharm Pharmacol Res 2024
DOI:10.26502/fjppr.0101

Learning rate	0.1
Limit depth of individual tress	3
Do not split subsets smaller than	2
Fraction of training instances	1
Tree	
Binary tree	
Minimum number of instances in leaves	2
Do not split subsets smaller than	5
Limit the maximal tree depth to	100
Classification	Stop when majority reaches 95%
KNN	
Number of neighbors	5
Metric	Euclidean
Weight	Uniform
Neural Network	
Neuros in Hidden layers	100
Activation	ReLu
Solver	Adam
Regularization	a=0.0001
Maximal number of interations	200
Random Forest	
Number of trees	500
Number of attibulates considered at each split	3
Do not split subsets smaller than	5

Table 2: **A)** Classification performance metrics for models classifying samples of all groups (control, smokers, and occasional smokers). Evaluated metrics were Area Under the receiver operating characteristic Curve (AUC), classification accuracy (CA), F1 score, precision, recall. **B)** Confusion matrix for the most accurate model (Neural Network) classifying samples of all groups (control, smokers, and occasional smokers). The percentage of corrected classified instances are showing in the diagonal of the matrix (correspondence beween predicted and true classes for each group). Following the sequence, the classficiation performance metrics for models classifying samples of each pair of groups: Control x Smokers **(C)**, Control x Occasional smokers **(D)**, and Occasional smokers x Smokers **(E)**. Evaluated metrics were Area Under the receiver operating characteristic Curve (AUC), classification accuracy (CA), F1 score, precision, recall. Confusion matrix for the most accurate models classifying samples of each pair of groups: Control x Smokers **(F),** Control x Occasional smokers **(G)**, and Occasional smokers x Smokers **(H)**. The percentage of corrected classified instances are showing in the diagonal of the matrix (correspondence between predicted and true classes for each group).

A)	**Model**	**AUC**	**CA**	**F1**	**Precision**	**Recall**
kNN		0.717	0.584	0.575	0.588	0.584
Tree		0.754	0.634	0.633	0.632	0.634
SVM		0.783	0.634	0.612	0.638	0.634
SGD		0.722	0.644	0.635	0.633	0.644
Random Forest		0.8	0.634	0.621	0.627	0.634
Neural Network		0.857	0.713	0.712	0.715	0.713
Naive Bayes		0.676	0.545	0.535	0.536	0.545
Logistic Regression		0.641	0.475	0.379	0.397	0.475
Gradient Boosting		0.819	0.663	0.657	0.668	0.663
AdaBoost		0.692	0.604	0.601	0.605	0.604

B)		**Predicted**		
		Control	**Occasional Smokers**	**Smokers**
True class	**Control**	72.70%	20.70%	10.70%
	Occasional Smokers	9.10%	65.50%	14.30%
	Smokers	18.20%	13.80%	75.00%

C)	**Model**	**AUC**	**CA**	**F1**	**Precision**	**Recall**

Nogueira MS, et al., J Pharm Pharmacol Res 2024
DOI:10.26502/fjppr.0101

Control x Smokers	kNN	0.768	0.72	0.714	0.722	0.72
	Tree	0.748	0.733	0.734	0.739	0.733
	SVM	0.839	0.787	0.785	0.787	0.787
	SGD	0.85	0.853	0.853	0.853	0.853
	Random Forest	0.819	0.76	0.759	0.759	0.76
	Neural Network	0.894	0.84	0.837	0.846	0.84
	Naïve Bayes	0.737	0.68	0.68	0.68	0.68
	Logistic Regression	0.657	0.56	0.424	0.535	0.56
	Gradient Boosting	0.827	0.76	0.759	0.759	0.76
	AdaBoost	0.76	0.76	0.761	0.762	0.76
D)	**Model**	**AUC**	**CA**	**F1**	**Precision**	**Recall**
Control x Occasional smokers	kNN	0.752	0.712	0.714	0.725	0.712
	Tree	0.592	0.576	0.576	0.576	0.576
	SVM	0.858	0.738	0.786	0.787	0.788
	SGD	0.805	0.803	0.804	0.809	0.803
	Random Forest	0.842	0.758	0.754	0.756	0.758
	Neural Network	0.85	0.773	0.775	0.785	0.773
	Naïve Bayes	0.57	0.515	0.519	0.537	0.515
	Logistic Regression	0.429	0.591	0.439	0.349	0.591
	Gradient Boosting	0.782	0.773	0.773	0.774	0.773
	AdaBoost	0.742	0.742	0.744	0.749	0.742
E)	**Model**	**AUC**	**CA**	**F1**	**Precision**	**Recall**
Occasional smokers x Smokers	kNN	0.704	0.738	0.738	0.753	0.738
	Tree	0.588	0.623	0.624	0.625	0.623
	SVM	0.797	0.738	0.738	0.738	0.738
	SGD	0.719	0.721	0.722	0.723	0.721
	Random Forest	0.818	0.754	0.754	0.755	0.754
	Neural Network	0.854	0.787	0.787	0.788	0.787
	Naïve Bayes	0.713	0.705	0.705	0.705	0.705
	Logistic Regression	0.563	0.557	0.399	0.311	0.557
	Gradient Boosting	0.794	0.721	0.722	0.727	0.721
	AdaBoost	0.738	0.738	0.738	0.741	0.738
F)					**Predicted**	
					Control	**Smokers**
True class		**Control**			86.00%	15.60%
		Smokers			14.00%	84.40%

Nogueira MS, et al., J Pharm Pharmacol Res 2024
DOI:10.26502/fjppr.0101

G)					Predicted	
					Control	Occasional Smokers
True class		Control			76.70%	26.10%
		Occasional Smokers			23.30%	73.90%
H)					Predicted	
					Occasional Smokers	Smokers
True class		Occasional Smokers			71.40%	21.20%
		Smokers			28.60%	78.80%

Discussion

For the following study, the biofluid of choice was saliva, due to its characteristics, such as being easily collected using a non-invasive technique. Several studies conducted with this biofluid have demonstrated its effectiveness for diagnosis, including the fact that researchers have demonstrated that the content of saliva alterations is closely related to the onset of oral diseases and systemic diseases [2, 12, 13, 24]. Our results suggested that biochemical changes can be observed mainly in DNA, Proteins, Collagen, Amide I and Amide II bands between saliva of the study groups (non-smokers, smokers, and occasional smokers). These changes may be associated with oral cancer, as smoking is a risk factor [2], [8]. In general, FT-IR spectroscopy was effective in differentiating these groups, especially between control and smokers. Classification models built with SGD and Neural Network methods were the most accurate.

Once determining which vibrational modes are similar between saliva FT-IR spectra of smokers and cancer patients, following up the presence of these vibrational modes in a larger population and over a longer period of time (e.g., 2 or 5 years follow up studies), this will allow us to investigate which modes and corresponding biochemical changes (biomarkers) are associated to the development of oral cancer by smoking as its main risk factor. At this stage, FT-IR spectroscopy can be used as a point-of-care (POC) diagnostics tool to identify oral cancer biomarkers, given that the FT-IR technology is sufficiently mature for compact and cost-effective instruments to be produced [2], [24]. These instrument production aspects combined with the small sample volume required and real-time sample analysis provided by FT-IR spectroscopy potentially enables quicker clinical translation and commercialization as fundamental steps towards implementation of POC technologies for high-throughput screening tests. It is important to note that all classification performance metrics have been obtained with leave-one-out validation, which illustrates the maximum performance the classification model will reach by assuming the biological variability of the dataset is the same as that of larger populations. Increasing the sample size over the numbers of our preliminary study will help to confirm whether the same metrics are achievable. Finally, our study is a pilot study in which the number of patients should be increased to evaluate features of different forms of smoking substances in FT-IR spectra of saliva samples of smokers, as well as find similarity of these spectra with FT-IR spectra of cancer patients. The technique was effective for characterizing the samples and could be used as a tool for analysis of saliva [2].Future steps include increase the number of volunteers involved in the study to incorporate more data on the biological variability into classification models and for subsequent exploratory analysis.

Conclusion

It is worth noting that a clinical translation requirement of FT-IR is the training of clinicians who can use it daily in the clinic. Therefore, early involvement of multiple clinical research centers and the implementation of robust machine learning methods providing scores of risk of oral cancer for individual patients even with preliminary datasets would be beneficial to engage clinicians in a growing multicenter study, potentially culminating in the development of technologies for cancer screening programs nationally and internationally. If these screening programs are implemented, benefits to patients range from early cancer detection to improved prognosis, while clinicians benefit from a larger number of patients complying to cancer screening, as FT-IR spectroscopy enables non-invasive and non-destructive analysis of saliva samples unlike conventional biopsies followed by histopathological analysis. In a long-term, initiating cancer treatments at earlier stages could make a difference to every patient independently on his/her socio-economic status.

Acknowledgements

Marcelo Saito Nogueira received his salary from Science Foundation Ireland (SFI/15/RP/2828) and SFI-22/RP-2TF/10293

Nogueira MS, et al., J Pharm Pharmacol Res 2024
DOI:10.26502/fjppr.0101

Funding: This study was not supported by any funding.

Conflict of Interest: The authors declare that they have no conflict of interest.

References

1. Zapata F, Fernández de la Ossa MÁ, and García-Ruiz C. 'Emerging spectrometric techniques for the forensic analysis of body fluids (2015).
2. Rodrigues LM, Magrini TD, Lima CF, Scholz J, da Silva Martinho H, et al. 'Effect of smoking cessation in saliva compounds by FTIR spectroscopy', Spectrochim Acta A Mol Biomol Spectrosc 174 (2017): 124–129.
3. Michcik A. et al. 'Tobacco smoking alters the number of oral epithelial cells with apoptotic features', Folia Histochem Cytobiol 52 (2014): 60–68.
4. Thun M, Peto R, Boreham J, and Lopez AD. 'Stages of the cigarette epidemic on entering its second century', Tob Control 21 (2012): 96–101.
5. Leal LB, Nogueira MS, Canevari RA and Carvalho LFCS. 'Vibration spectroscopy and body biofluids: Literature review for clinical applications', Photodiagnosis Photodyn Ther 24 (2018): 237–244.
6. Hoffmann D, Hoffmann I and El-Bayoumy K. 'The less harmful cigarette: A controversial issue. A tribute to Ernst L. Wynder' (2001).
7. Takahashi K, ichi Yokota S, Tatsumi N, Fukami T, et al. 'Cigarette smoking substantially alters plasma microRNA profiles in healthy subjects', Toxicol Appl Pharmacol 272 (2013): 154–160.
8. Musk AW and De Klerk NH. 'History of tobacco and health', Respirology 8 (2003): 286–290.
9. Vista do Patologias bucais relacionadas ao tabagismo'. Accessed: Aug 25 (2024).
10. Cabral AR, et al., 'Os Impactos negativos do uso do cigarro eletrônico na saúde', Diversitas Journal 7 (2022): 0277–0289.
11. Goulart D, et al., 'Tabagismo em idosos Smoking in the elderly' (2006).
12. Castagnola M, PiCCiotti P, Fanali C, Passali GC, and iavarone F. 'Potential applications of human saliva as diagnostic fluid Le potenziali applicazioni della saliva umana come fluido diagnostico', ACTA oTorhinolAryngologiCA iTAliCA 31 (2011): 347–357.
13. Y-H Lee, DT Wong, D Felix, M Yip. Endowed Professor, and A. Dean, 'Saliva: An emerging biofluid for early detection of diseases'
14. MÉTODO ÓPTICO PARA DIAGNÓSTICO DE CÂNCER ORAL VIA SALIVA; USO DO MÉTODO E PROTÓTIPO'.
15. Mendes De Freitas R. Maria A, Rodrigues X, Ferreira De Matos Júnior A, et al. 'Fatores de risco e principais alterações citopatológicas do câncer bucal: uma revisão de literatura Risk factors and major cytopathologycal changes of oral cancer: a review of literature' (2016).
16. Das LF, De Carvalho CeS and Saito Nogueira M. 'New insights of Raman spectroscopy for oral clinical applications', Analyst 143 (2018): 6037–6048.
17. Cosci A, et al., 'Time-resolved fluorescence spectroscopy for clinical diagnosis of actinic cheilitis: erratum' (2018).
18. Felipe L, Carvalho CS, Saito Nogueira M, Neto LPM, Bhattacharjee TT, and Martin AA. 'Raman spectral post-processing for oral tissue discrimination – a step for an automatized diagnostic system', Biomed Opt Express 8 (2017): 5218–5227.
19. Naseer K, Ali S and Qazi J. 'ATR-FTIR spectroscopy as the future of diagnostics: a systematic review of the approach using bio-fluids. (2021).
20. Bunaciu AA, Hoang VD, and Aboul-Enein HY. 'Applications of FT-IR Spectrophotometry in Cancer Diagnostics', Crit Rev Anal Chem 45 (2015): 156–165.
21. Cesar P, Júnior C, Ferreira Strixino J and Raniero L. 'Analysis of saliva by Fourier transform infrared spectroscopy for diagnosis of physiological stress in athletes'.
22. Shrestha B and Dunn L. 'The Declaration of Helsinki on Medical Research involving Human Subjects: A Review of Seventh Revision', J Nepal Health Res Counc 17 (2020): 548–552.
23. Movasaghi Z, Rehman S and Rehman IU. 'Fourier Transform Infrared (FTIR) Spectroscopy of Biological Tissues', Appl Spectrosc Rev 43 (2008): 134–179.
24. Zhang A, Sun H, and Wang X. 'Saliva metabolomics opens door to biomarker discovery, disease diagnosis, and treatment', Appl Biochem Biotechnol 168 (2012): 1718–1727.