matters

A low resource setting oral lesion screening tool

Pekka E Hänninen, Joonas J Siivonen, Eija Martikkala, Diti Desai, Timo Teimonen, Satu A Tiittanen, Paul Mundill, Chetan Mukundan, Rani Desai, Praveen Birur

Laboratory of Biophysics, Institute of Biomedicine, Faculty of Medicine, University of Turku; R&D, Aqsens Inc.; Main office, Aqsens Inc; R&D, Aqsens Inc; Axxonet Bangalore, Axxonet System Technologies Pvt Ltd; Biocon Foundation, Biocon Foundation; Dental sciences, K.L.E.S Institute of Dental sciences

Triple Blind Peer Review
The handling editor, the reviewers, and the authors are all blinded during the review process.

Full Open Access
Supported by the Velux Foundation, the University of Zurich, and the EPFL School of Life Sciences.

Abstract

We demonstrate the use of saline mouth rinse and chemical fingerprinting approach in screening for oral lesions from tobacco product using patients in low resource settings in rural India. The method is based on the use of a nonspecific long lifetime luminescent lanthanide label, time-resolved measurement and modulator chemistry that is sensitive to ions, proteins and pH but without specific probes. The sensitivity and specificity of the method bases on the use of trained classification algorithms. The overall lesion detection accuracy of the method is shown to be about 90% with the tested patient material utilizing two differently trained classification algorithms.

Introduction

Data from the recent Million Death Study in India [1] shows that cancer of oral cavity (including lip and pharynx) was one of most fatal cancers in males and accounts for 35% of all deaths due to cancer. The main attributing reason for the high prevalence of oral cavity cancer in India and southeast Asia is the high tobacco product usage. Due to its asymptomatic onset, over 60% of the patients are diagnosed after an advanced stage of disease [2]. Patients often delay or may never follow up to see the specialists for further treatment. Under such circumstances, increasing cancer treatment resources will have very little impact on the outcome. The most cost-effective means to impact cancer outcome in the community is to channel resources to prevention of preventable cancers, and early detection and down-staging of cancer and provide standardized protocol-based treatment of diagnosed patients. Thus simple, easy-to-use and cost-effective tools utilizing, for example, salivary sampling for risk assessment and early detection of oral cavity cancer would be required.

The use of saliva for diagnostics has been studied extensively because of its noninvasive nature and potential applicability to low resource settings. There are, however, serious drawbacks related to salivary sampling and its constitution: individual variance of sampling is always very large depending on several factors from salivation to personal habits affecting the constitution of saliva and potentially masking specific markers; also the concentrations of the target markers vary. Literature, however, suggests that there are several markers in saliva that could be used for aiding the diagnosis of oral cancer. Markopoulos et al. [3] summarize these findings in their review article stating that "saliva seems to emerge as a valuable tool in cancer diagnostics and mass population screening." Zhang et. al take these expectation further in their very recent review [4], but these expectations, especially in view of low resource settings, are yet largely to realize.

Of practical approaches, observation of Gene Promoter Hypermethylation by Methylation Specific PCR (MSP) from unstimulated drool saliva has recently been shown to be a promising approach with more than 80% overall accuracy [5]. Pereira and Franzmann et. al [6] [7] have studied the use of markers CD44, IL-8, HA and total protein concentration from oral rinses; by combining these markers with data of personal risk factors they could reach an accuracy of 70–80% depending of the selected multivariate analysis parameter cut-off values. Whereas DNA amplification based methods present the ultimate sensitivity in detection of a specific target, they are not at the current stand of technology convertible to low resource settings but require a specialized laboratory and personnel. Results from studies of Pereira and Franzmann, however, suggest that a point-of-care, simple and low-cost test strip could be developed. However, in low resource settings no patient follow-up records exist and questionnaires made at time of meeting the health care worker are unreliable so the test should not rely on data on personal habits but objectively classify the condition of the patient.

Objective

Our aim was to study an approach where oral rinse samples from tobacco product using patients were measured with a simple, fuzzy, nonspecific array and the data analyzed with multivariate algorithms to classify the samples to two groups: healthy and lesion. The system is simple, low cost and transferable to low resource compatible form- the question whether it works or not for the intended screening purpose of oral lesions leading to cancer needed to be solved.
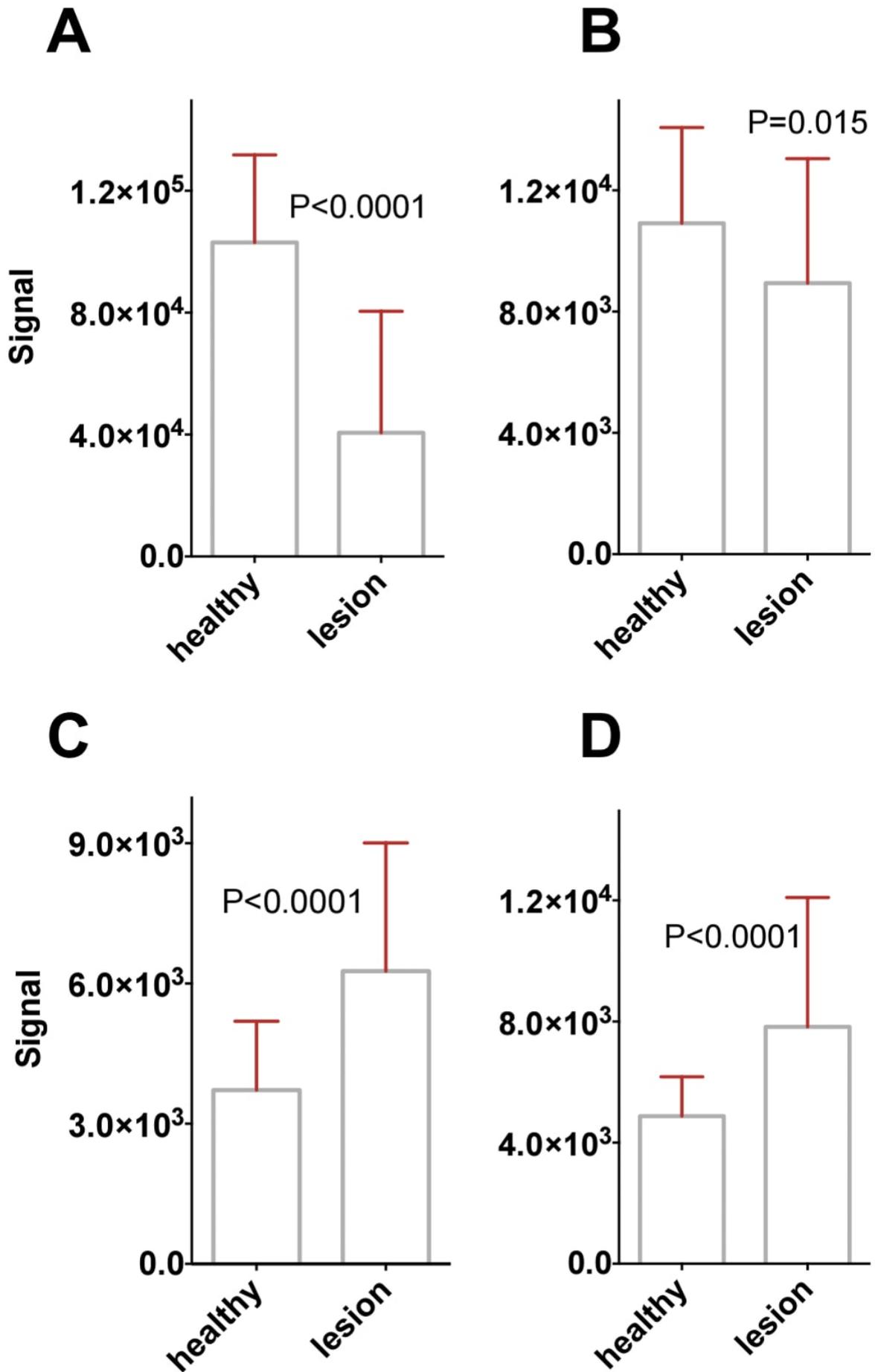
Figure 1

Figure Legend

Figure 1. Data distribution and separation by four different semi-specific chemistries.

The P-values of the t-test are given for unpaired t-test assuming independent SD for each population. The letters A through D refer to the used chemistries from the list given in "Methods" section. The error bars are STDEV of the measured data.

As the sample we used an oral rinse. Each of the patients rinsed their mouth with 15 ml of 0.9% NaCl solution which was collected and immediately stored in ice and after transport to the lab in Axxonet, Bangalore at -20℃ (collection protocol modified from Rettori et. al. 11). The number of patients in this study was 53 with 40 patients presenting lesions and 13 healthy patients (see also "Limitations" section). The patients were all working age adults between 20 and 60 years of age and of both sexes.

The patient status was determined in a visual examination by a clinician and recorded along with personal information and information of the tobacco use. The lesions in the tested data set were found in either buccal or lower labial mucosa and no distinction in the teaching of the method was made between the positions of the lesions.

The measurement protocol was as follows

1. The sample from the freezer is thawed, and Eu-terpyridine chelate (Ramidus Ab, Lund, Sweden) is added to concentration of 0.76 nM.

2. 70 µl sample is dispensed in duplicate to a microtiter plate containing the reagents dispensed in water (see the reagent list below).

3. Incubate 10 min at room temperature.

4. Measure with TRF plate-reader (Labrox Inc. Turku, Finland) with 340 nm excitation and 615 nm emission filters with delay time of 400 ms and measurement window of 400 ms. Each measurement was 100 cycles (flashes).

The used nonspecific chemistries in the wells were

- A: o-cresolphtalein complexone (1006 µM final concentration), pH sensitive chemistry.
- B: EGTA (50 µM) + saffrosine (52 µM), protein sensitive chemistry.
- C: EGTA (50 µM) + pyrocathecol violet (101 µM), ion/complex sensitive chemistry.
- D: 3-HPA (50 µM) + bromocresol purple (71 µM), pH & albumin sensitive chemistry.

The measured data was first analyzed for statistical differences between the two groups (healthy, lesion) for each of the chemistries A-D. An unpaired parametric t-test was calculated using Prism (version 6) software with the assumption that the SDs for the populations are not the same. The second step of analysis was to generate the training and test sets of the data with Molegro data modeler version 2.1. Since the amount of data was rather limited with 13 healthy and 40 lesion patients, we decided to generate an artificial train set by taking the average signal of each group and chemistry as the starting point. A training set of 300 samples was then generated by adding randomly distributed noise to the average signals. The standard deviation of the noise was the same as that of the real data and this artificial noise data was evenly distributed around the mean of 0. The test sets were the measured raw data set with the replicates as individual points and another one with the replicates averaged. The used classification methods from Molegro were K-Nearest Neighbor (KNN) and Support Vector Classification (SVC). The fine-tuning function of the program was used in finding the parameters for optimal training. We also tried splitting the measured data randomly into individual training and testing sets, but due to the low amount of data the results varied more than with the adopted method depending on the split (data not shown). **Results &**

Discussion

The patient samples were collected as a part of the oral cancer screening program in India and were measured with the newly developed 4-space array utilizing environmentally sensitive lanthanide chelate chemistry and time-resolved luminescence (TRL) measurement in conjunction with indicator dye molecules and modulating (buffer) chemistries. The rationale of using the inherently environmentally sensitive lanthanide chemistry along with indicator molecules is to tune each of the 4-space TRL signals to

give a unique response that is dependent on the interaction of the indicator molecules, the sample and the modulator chemistry with the luminescent lanthanide. Since the interactions are nonspecific, each of the 4-space signals is indicative of a broader property of the mouth rinse samples than presence of any single marker molecule. When the individual 4-space measurements are sufficiently independent, we claim that their joint information can become specific in view of the question at hand. Observing the absolute TRL signals from figure 1, along with the behavior of these signals with respect to the two groups to be separated, we claim that the used chemistry combinations give sufficiently uncorrelated responses for each of the 4-space data to support good classification.

The used data classification methods were K-Nearest Neighbor (KNN) and Support Vector Classification (SVC). The fine-tuning function of the data analysis program (Molegro Data Modeler 2.1) was used in finding the parameters for optimal training. The results for the data analysis are given in figure 1, and are listed below. The t-test indicated that each of the selected chemistries alone showed significant difference ($P<0.05$) between the healthy and lesion patient groups. From the tested two algorithms both performed equally well yielding sensitivity and specificity near the 90% mark. The use of an artificially and randomly generated training set from the patient ensemble averages ensured that we were able to assess all our patient data and that the algorithms could not be over-trained to recognize false linkages between the samples.

The results from the analysis for detection of lesions are as follows:
- KNN (k=2), all data: sensitivity/specificity - 89/88 %
- SVC, all data: sensitivity/specificity - 90/88%
- KNN, averaged replicates: sensitivity/specificity - 88/92%
- SVC, averaged replicates: sensitivity/specificity - 90/92%

As can be seen, the prediction sensitivity and specificity approach the 90% mark regardless of the classification method. Averaging or alternatively using the replicates as individual measurements did not have a significant effect on the results, suggesting that the actual measurements were sufficiently repeatable for the algorithms to function well. This was also supported by the observation of the variance displayed in figure 1.

In all, results indicate that nonspecific means can be used in differentiating saliva samples by using multivariate analysis tools. In the collection of data, the prerequisite (inclusion criterion) for measurement was that the person was a tobacco product user. By concentrating to the most likely population at risk, we could train our analysis to detect differences between tobacco users and did not experience the well-known problem of tobacco product interference in salivary diagnostics in the training phase of the system. Further, in a learning-based self-calibrating system, the use of standards and controls is only necessary to verify the instrument, the performance of the algorithms and the method relies on the information from a trained clinician at the teaching phase. In fact, considering the function of the method, the use of, for example, pooled samples or artificial samples does not contribute to testing of the function of the method since we are profiling the sample rather than picking certain features; artificial samples are, against our normal view, obscure in this respect.

Conclusions

Our data shows that chemical fingerprinting could be an alternative to assessment of oral lesions. Chemical fingerprinting is an interesting approach for low resource setting diagnostics since the chemistries are simple and cost effective to reproduce and the measurement protocol is simple. Although our current data set did not contain any detectable early malignant lesions, we believe, that the results are indicative that such assessment tool can be developed utilizing our "fuzzy" chemical fingerprinting.

Limitations

The study was performed as part of a larger oral cancer screening program where the capability of healthcare workers to assess oral lesions was tested along with a mobile solution to transfer images of the most severe lesions to specialists. The sample collection was not originally planned as part of the program and turned out to be challenging, es-

pecially in terms of performing biopsy from suspected oral cancer cases. The number of biopsies remained thus low and none of them turned out positive early stage cancer, so this study will require a follow-up. In all, more than 400 patient samples were collected during the process of developing the chemistry with the converged results showed in this paper for the final remaining samples after the development work on-site. The samples were also not further analyzed by other means (e.g. "omics", HPLC), leaving exact explanations of the sample differences out of the scope of this work.

The study shows that the fingerprinting approach is a viable option for saliva diagnostics. In general, the use of multivariate methods seems to be an emerging trend in the most difficult areas of diagnostics, and our work follows this track, although the approach is more holistic than usual. However, to confirm our findings and to take the method nearer to practice more data will be needed, in particular of early cases of oral cancer.

Additional Information

Methods

As the sample we used an oral rinse. Each of the patients rinsed their mouth with 15 ml of 0.9% NaCl solution which was collected and immediately stored in ice and after transport to the lab in Axxonet, Bangalore at -20℃ (collection protocol modified from Rettori et. al. 11). The number of patients in this study was 53 with 40 patients presenting lesions and 13 healthy patients (see also "Limitations" section). The patients were all working age adults between 20 and 60 years of age and of both sexes.

The patient status was determined in a visual examination by a clinician and recorded along with personal information and information of the tobacco use. The lesions in the tested data set were found in either buccal or lower labial mucosa and no distinction in the teaching of the method was made between the positions of the lesions.

The measurement protocol was as follows

1. The sample from the freezer is thawed, and Eu-terpyridine chelate (Ramidus Ab, Lund, Sweden) is added to concentration of 0.76 nM.

2. 70 µl sample is dispensed in duplicate to a microtiter plate containing the reagents dispensed in water (see the reagent list below).

3. Incubate 10 min at room temperature.

4. Measure with TRF plate-reader (Labrox Inc. Turku, Finland) with 340 nm excitation and 615 nm emission filters with delay time of 400 ms and measurement window of 400 ms. Each measurement was 100 cycles (flashes).

The used nonspecific chemistries in the wells were

- A: o-cresolphtalein complexone (1006 µM final concentration), pH sensitive chemistry.
- B: EGTA (50 µM) + saffrosine (52 µM), protein sensitive chemistry.
- C: EGTA (50 µM) + pyrocathecol violet (101 µM), ion/complex sensitive chemistry.
- D: 3-HPA (50 µM) + bromocresol purple (71 µM), pH & albumin sensitive chemistry.

The measured data was first analyzed for statistical differences between the two groups (healthy, lesion) for each of the chemistries A-D. An unpaired parametric t-test was calculated using Prism (version 6) software with the assumption that the SDs for the populations are not the same. The second step of analysis was to generate the training and test sets of the data with Molegro data modeler version 2.1. Since the amount of data was rather limited with 13 healthy and 40 lesion patients, we decided to generate an artificial train set by taking the average signal of each group and chemistry as the starting point. A training set of 300 samples was then generated by adding randomly distributed noise to the average signals. The standard deviation of the noise was the same as that of the real data and this artificial noise data was evenly distributed around the mean of 0. The test sets were the measured raw data set with the replicates as individual points and another one with the replicates averaged. The used classification methods from Molegro were K-Nearest Neighbor (KNN) and Support Vector Classification (SVC). The fine-tuning function of the program was used in finding the parameters for optimal training. We also tried splitting the measured data randomly into individual training and testing sets, but due to the low amount of data the results varied more than with the adopted method depending on the split (data not shown).

Supplementary Material

Please see https://sciencematters.io/articles/201603000002.

Ethics Statement

The collection was performed under institutional ethical permit of Biocon Foundation. The institutional ethical committee consented for the project. (Ethical clearance: KIDS/IEC/11-2014/30). An informed consent was obtained from all the participants.

Citations

[1] Rajesh Dikshit et al. "Cancer mortality in India: a nationally representative survey". In: *The Lancet* 379.9828 (May 2012), pp. 1807–1816. DOI: 10.1016/s0140-6736(12)60358-4. URL: http://dx.doi.org/10.1016/S0140-6736(12)60358-4.

[2] Gonsalves et al. "Common Oral Lesions: Part II. Masses and Neoplasia." In: *American Family Physician* 75 (2007), pp. 509–12.

[3] Anastasios K. Markopoulos, Evangelia Z. Michailidou, and Georgios Tzimagiorgis. "Salivary Markers for Oral Cancer Detection". In: *The Open Dentistry Journal* 4.1 (Sept. 2010), pp. 172–178. DOI: 10.2174/1874210601004010172. URL: http://dx.doi.org/10.2174/1874210601004010172.

[4] Yong Zhang et al. "The emerging landscape of salivary diagnostics". In: *Periodontology 2000* 70.1 (Dec. 2015), pp. 38–52.

DOI: 10.1111/prd.12099. URL: http://dx.doi.org/10.1111/prd.12099.

[5] Dmitry A. Ovchinnikov et al. "Tumor-suppressor Gene Promoter Hypermethylation in Saliva of Head and Neck Cancer Patients". In: *Translational Oncology* 5.5 (Oct. 2012), pp. 321–326. DOI: 10.1593/tlo.12232. URL: http://dx.doi.org/10.1593/tlo.12232.

[6] Pereira et al. "Salivary markers and risk factor data: a multivariate modeling approach for head and neck squamous cell carcinoma detection." In: *Cancer biomarkers : section A of Disease markers* 10 (2011), pp. 241–9.

[7] Elizabeth J. Franzmann et al. "Salivary protein and solCD44 levels as a potential screening tool for early detection of head and neck squamous cell carcinoma". In: *Head and Neck* 34.5 (July 2011), pp. 687–695. DOI: 10.1002/hed.21810. URL: http://dx.doi.org/10.1002/hed.21810.