Feasibility of Desorption Electrospray Ionization Mass Spectrometry for Diagnosis of Oral Tongue Squamous Cell Carcinoma

Authors: Cedric D'Hue*, Michael Moore**, Don-John Summerlin***, Alan Jarmusch*, Clint Alfaro*, Avinash Mantravadi****, Arnaud Bewley**, D. Gregory Farwell**, R. Graham Cooks*+ —

* Department of Chemistry, Purdue University,
560 Oval Drive
West Lafayette, Indiana 47907-2084

** University of California at Davis, Comprehensive Cancer Center
Department of Otolaryngology-Head and Neck Surgery
2521 Stockton Blvd., Suite 7200
Sacramento, CA 95817

***Indiana University Health Pathology Laboratory
350 West 11th Street, Room 4022
Indianapolis, IN 46202-4108

****Indiana University Department of Otolaryngology-Head and Neck Surgery
550 N. University Blvd. Rm 3170
Indianapolis, IN 46202

Corresponding author: R. Graham Cooks, cooks@purdue.edu

Abstract

Rationale: Desorption electrospray ionization-mass spectrometry (DESI-MS) has demonstrated utility in differentiating tumor from adjacent normal tissue in both urologic and neurosurgical specimens. We sought to evaluate if this technique had similar accuracy in differentiating oral tongue squamous cell carcinoma (SCC) from adjacent normal epithelium due to current issues with late diagnosis of SCC in advanced stages.

Methods: Fresh frozen samples of SCC and adjacent normal tissue were obtained by surgical resection. Resections were analyzed using DESI-MS sometimes by a blinded technologist. Normative spectra were obtained for separate regions containing SCC or adjacent normal epithelium. Principal Component Analysis and Linear Discriminant Analysis (PCA-LDA) of spectra were used to predict SCC versus normal tongue epithelium. Predictions were compared with pathology to assess accuracy in differentiating oral SCC from adjacent normal tissue.

Results: Initial PCA score and loading plots showed clear separation of SCC and normal epithelial tissue using DESI-MS. PCA-LDA resulted in accuracy rates of 95% for SCC versus normal and 93% for SCC, adjacent normal and normal. Additional samples were blindly analyzed with PCA-LDA pixel by pixel predicted classifications as SCC or normal tongue epithelial tissue and compared against histopathology. The *m/z* 700-900 prediction model showed a 91% accuracy rate.

Conclusion: DESI-MS accurately differentiated oral SCC from adjacent normal epithelium. Classification of all typical tissue types and pixel predictions with additional classifications should increase confidence in the validation model.

Introduction

Oral Cavity Cancer

Cancers of the oral cavity and oropharynx are the most common head and neck tumors in the United States.[1] Worldwide, around 600,000 people are afflicted by head and neck cancer every year.[2] Oral cancer has an age normalized incidence rate of 3.9 per 100,000 persons worldwide.[3] Common risk factors for developing oral cancer are alcohol consumption and tobacco use with a rising incidence of oropharyngeal cancer (more specifically, tonsil and base of tongue cancer) related to human papilloma virus (HPV) infection.[3]

Oral cancer patients typically are diagnosed at an advanced stage of invasive squamous cell carcinoma (SCC) at the time of diagnosis.[4] Owing to delayed disease presentation in patients and lack of appropriate screening techniques to identify early oral cancer, roughly half of patients survive five years from initial disease diagnosis.[2] The five year oral cancer mortality rate has not changed greatly in over 50 years[4] despite surgical advances, and improvements in radiotherapy and chemotherapy. Oral cancer not only has a very high recurrence rate, but those who survive a first occurrence also have up to a 20 fold increased risk of developing a second primary cancer[3] (i.e., another oral SCC) which is the highest second-malignancy rate for all cancers.[4] The most common cause of treatment failure and death in oral cancer patients is local recurrence of their primary tumor.[5]

The high frequency of patients presenting with advanced disease is puzzling. Oral cancer is located in the oral cavity where the oral surface epithelium is readily accessible to visual and tactile examination. Health care providers, specifically dentists, typically examine their patient's oral cavity on a semi-annual basis providing a consistent and reliable opportunity for a visible and palpable examination of areas at risk.

Current Oral Cavity Screening methods

Current methods in oral cavity cancer screening (*e.g.* visible and palpable examination of oral cavity by a dentist, brush tests, toluidine blue staining, and light-based detection systems) do not seem to definitively determine the presence of cancer or precancerous lesions.[6] Conventional oral examination (COE) is currently a controversial mainstay of oral cancer screening.[4,7] Brush tests act more as a case-finding modality by interrogating clinical lesions that would not typically undergo tissue biopsy and pathological evaluation. However, since not all subjects interrogated by brush test undergo tissue biopsy and pathological evaluation, their accuracy is still unclear.[4]

One study evaluated the use of toluidine blue to detect lesions that had not been detected by COE[8] concluding that the staining test was sensitive (92%) but not specific (42%) since there is no consensus on what shade of blue is a positive result.[4]

The lack of comparison of light-based detection system test results to pathological evaluation makes sensitivity, specificity and similar calculations impossible.[4] A case-finding study of 44 patients with a history of oral dysplasia or head and neck SCC (HNSCC) evaluated by tissue fluorescence resulted in a 98% sensitivity and 100% specificity in discriminating cancer or dysplasia from normal oral mucosa.[9]

Currently the only definitive method of detection is by biopsy and pathological evaluation of removed tissue.[10,11] However, even in this circumstance, while evaluation by the pathologist is directed by numerous criterion for diagnosis, it still is inherently subjective.

Desorption Electrospray Mass Spectrometry as a diagnostic method

Subjective histopathological analysis would likely be improved by objectively relating pathological states to chemical information obtained by desorption electrospray ionization – mass spectrometry (DESI-MS).[12,3,13,1]

As established in several other organ systems, DESI-MS imaging of lipid profiles in tissue can be correlated with histopathological characterization.[14] DESI-MS benefits from minimal sample handling beyond what is needed for frozen section pathological evaluation making DESI-MS well suited for case finding. Previous work in brain surgery[15] has differentiated types of gliomas and meningiomas, including tumors of different histological grades and tumor cell concentrations, and was validated during intraoperative use[16].

There is an urgent need for earlier objective detection of oral cancer for diagnosis and prognosis. A goal of this project is to develop a new diagnostic aid that could help a health care professional, such as a general dentist or dental specialist, more readily identify and assess oral tissue of uncertain biological significance, defined as case-finding.[4]

Methods and Materials

Tissue Specimens

IUSOM and UC Davis (UCD) research tissue specimens were collected during tumor removal surgeries conducted by Dr. Michael G. Moore, MD, FACS (otolaryngologist, previously at IUSOM and currently at UC Davis Medical Center) and other surgeons in the Department of Otolaryngology-Head and Neck Surgery at Indiana University and at UC Davis Medical Center. Each institution collected fresh frozen normal and SCC tissue specimens from the oral cavity of consented patients undergoing treatment following an IU IRB approved procurement protocol. This study was carried out in strict accordance with the UCD, IU and Purdue IRB protocols.

Most specimens were matched, meaning that both normal and SCC specimens were procured from the same patient. Thirty-two normal and SCC specimens were acquired from an IUSOM tissue bank. Sixty-two normal and SCC specimens were acquired from UC Davis.

All specimens were shipped on dry ice to Indiana University Health Pathology Lab (IUHPL) where the specimens were cryosectioned to 15 µm thickness and thaw mounted to glass microscope slides. Tissue-Tek O.C.T. (Sakura Finetek) was used as the embedding agent. Regarding thaw mounting, typically two tissue sections were thaw mounted per slide and multiple slides were prepared per specimen. Sections were stored at -80°C prior to ambient ionization MS analysis. -80° C sections were vacuum dried for 15 minutes prior to DESI-MS imaging analysis.

DESI-MS imaging

A laboratory-built DESI ion source was coupled to a linear ion trap mass spectrometer (LTQ) controlled by Xcalibur 2.0 software (ThermoFisher Scientific, Waltham, MA). The LTQ MS and Xcalibur software were used in all experiments. The embedding agent used for tissue cryosectioning produces significant background signals in the positive ion mode but not in the negative ionization mode. For this reason, we chose to use negative ionization for all experiments. For DESI experiments, the automatic gain control (AGC) was inactivated. Each tissue section selected for analysis was DESI-MS imaging analyzed using morphologically friendly spray solvent,[17] dimethylformamide-acetonitrile (1:1) solution. Both solvents were purchased from Mallinckrodt Baker Inc. (Phillipsburg, NJ) and delivered at a flow rate of 2.0 µL min^{-1} using the instrument syringe pump. The DESI source parameters were set as follows:

capillary temperature 275 °C, voltage applied to the stainless steel needle syringe 5 kV, capillary voltage -25 V, tube lens voltage -115 V, capillary incident angle 52°, spray to surface distance ~3 mm, sample to inlet distance ~5 mm, and nitrogen gas at 180 psi. The LTQ was set to an injection time of 250 µs and two microscans. Tissue sections were analyzed using a moving stage with a lateral ("x" axis) scan rate of 307.69 µm s^{-1} in horizontal rows separated by a 200 µm vertical step ("y" axis).[18] Instrument scan time was coordinated with scan speed providing ~200 x 200 µm pixels. Full scan mass spectra were acquired in negative ion mode in the mass range m/z 200-1000 for each tissue section.

Pathology

The same slide analyzed by DESI-MS imaging was hematoxylin & eosin (H&E) stained, photographed, cover slip mounted, and histopathologically evaluated by an expert pathologist, (DJS). The sections were evaluated to identify regions of normal epithelium and SCC. Other tissue types such as abnormal, premalignant, or precancerous epithelium, skeletal and smooth muscle, fat, connective tissue, blood vessels, and the like were observed and noted during histological evaluation. As illustrated by a tissue section in Figure 1A and B, normal epithelial and squamous cell carcinoma regions of interest were delineated by line annotations.

Twelve normal and SCC IUSOM specimens were excluded from further analysis because the evaluated sections did not include either regions of normal epithelium or SCC. Sixteen of the sixty-two specimens acquired from UCD were excluded from further analysis because the evaluated sections did not include either regions of normal epithelium or SCC. All histological evaluations were provided only after DESI-MS imaging and pixel prediction model analysis. The researchers conducting DESI-MS imaging and pixel prediction model analysis were blinded to initial diagnosis information.

BioMap for DESI-MS Imaging selection of Pixels of Interest

BioMap (Novartis Institutes for BioMedical Research, Basel Switzerland)[18] was used to correlate histologically evaluated annotations to identify DESI-MS image pixels of interest. An in-house data conversion software, ImgConverter v3.0, was used to compile and convert Xcalibur MS data files (*.raw) into BioMap compatible ANALYZE format files (*.img).[19] BioMap allows visualization of the mass spectrum associated with each individual, "x", "y" position as well as mass spectra processing (e.g., mean and background subtraction) and selection of a region of interest by selection of x,y positions. Each x is a mass spectrum acquired along y horizontal row. Each x,y position is defined as a pixel.

All BioMap pixels assignments are based on histological evaluations provided by Dr. Summerlin independent of DESI analysis and imaging. Pixels of interest were highlighted conservatively, typically excluding tissue adjacent to line annotations, even to the exclusion of normal epithelial tissue from the pixels of interest.

One matched case of samples was excluded from DESI-MS imaging because normal epithelial could not be distinguished from connective for IUSOM Sample 30 and SCC tissue could not be distinguished from surface epithelial tissue for Sample 31. Specifically, BioMap m/z values could not visualize the tissue section in order to correlate BioMap pixels with histopathological evaluation.

Histology based regions of interest are translated into histologically guided selection of pixels of interest. Average mass spectra were obtained by averaging regions of interest.

Figures 1C and 1D display average mass spectra for squamous cell carcinoma and normal epithelial regions of interest, respectively, in Sample 35.

DESI-MS Image Data Processing

A data conversion software, ProteoWizard specifically MSConvertGUI, was used to convert Xcalibur MS data files (*.raw) into mzXML files (*.mzxml).[19]

Multivariate Statistical Analysis

Twenty-one DESI-MS averaged mass spectra from twenty separate IUSOM samples were used to create normal and SCC mass spectral test set databases. For DESI averaged mass spectra, regions annotated by pathology as normal epithelium or SCC with no other characterization were selected as groups of pixels of interest composed of the mass spectra at each pixel. Again, average mass spectra were obtained by averaging regions of interest.

Principal Component Analysis (PCA)

BioMap exported an average mass spectrum from the pixels of interest. Each average mass spectrum's list of m/z values and ion abundances was imported into Excel (Microsoft, Spokane, WA USA). Excel spreadsheets were used to import into Matlab (The MathWorks, Inc., Natick, MA USA). In-house scripts were used to transform each average mass spectrum by standard normal variate (SNV). All spectra for each DESI dataset were mean-centered for the same *m/z* value across all samples. PCA was performed on total ion count (TIC) normalized or SNV transformed, mean-centered data for DESI datasets.

Principal Component Analysis - Linear Discriminant Analysis (PCA-LDA)

Models were built using all of the objects based on each analyzed target set. Linear Discriminant Analysis (LDA) was applied on the DESI target datasets of SNV transformed, column centered mass spectra after compression by PCA, thereby using as variables the principal components instead of the original mass spectral data. LDA was performed as a supervised discriminant classification technique. Discriminant methods look for a delimiter that divides the global domain into a number of regions, each assigned to one of the classes. This delimiter identifies an open region for each class and such regions determine the assignment of the samples to one of the classes.[10] Model validation, specifically evaluation of the predictive ability of the model, was performed by means of cross-validation (CV).[10] For this study, five cross-validation deletion groups were selected, meaning that all of the samples (DESI dataset n=21) were divided five times systematically in a training set and a test set, with all of the samples being in the test set only once. A training set means objects were used for building the classification model. A test set means non-training set objects were used to evaluate the predictive ability of the model.

Pixel predictions and cross validations were used to estimate the sensitivity, specificity, and global prediction rate for all classes, as the CV prediction rate for each class, known as the complete validation strategy.[10] For CV prediction rate, sensitivity measures the proportion of subjects with the disease who test positive, while the specificity determines the proportion without the disease who test negative. For pixel prediction, sensitivity measures the proportion of tissue sections with squamous cell carcinoma, where the majority of the pixel predictions in the same region of tissue section were predicted squamous cell carcinoma, while the specificity determines the proportion of tissue sections with normal epithelial tissue where the majority of the pixel predictions in the same region of tissue section were predicted as normal epithelial tissue.

The predictive values determine the proportion of subjects with positive and negative test results that either do or do not have the disease. For pixel prediction, the prediction rate is the percentage of correct predictions on the objects in the test set relative to pathology. The CV prediction rate is the percentage of correct predictions on the objects in the CV test set. The confusion matrix shows how many samples belong to a certain category were correctly/incorrectly assigned by the classification rule to that category. In each matrix, each element gives the number of samples of the row category assigned to the column category. When a matrix includes only entries on the diagonal (e.g., there are no entries outside the main diagonal) there is perfect prediction of all the samples.

LDA models used for supervised discriminant classification were built using a range of two to six PCs. The models were chosen which provided the least number of false results (e.g., highest prediction rates) in cross validation.

Univariate Statistical Analysis

One way analysis of variance (ANOVA) was performed on the TIC normalized intensities for a subset of the detected ions using Origin Pro 2017 (OriginLab, Northampton, MA). The reported p-values result from the F-test comparing the mean TIC normalized ion intensities between the normal oral epithelial tissue samples and the oral SCC samples. P-values much less than 0.05 indicate that there is a significant difference between the two experimental groups for the specified *m/z* value. Statistical power analysis, a description of the probability of detecting a difference in means between experimental groups, was performed using JMP (SAS Institute, Cary, NC). Power ranges from 0 to 1; a power of 1 means that the probability of detecting an effect is 100% given that there is an effect to be detected.

Results

Three mass ranges of 200-400, 400-600, and 700-900 were used for data analysis of each DESI image. Two mass ranges (m/z 200-1,000 and 700-900) were used for pixel by pixel prediction models. Mass range 200-400 limits biochemical information to fatty acids, mass range 400-600 to fatty acid dimers, and 700-900 to complex phospholipids. All samples were analyzed using DESI-MS imaging over three consecutive days.

DESI-MS Imaging on Oral Tongue Cancer Samples

Regions of interest were correlated with histologically evaluated annotations by using specific *m/z* values. For example, Sample 35 histological evaluation identified a region of normal epithelial tissue as illustrated in the top left portion of Figure 1A, above. BioMap images of Sample 35 at *m/z* 465 (cholesterol sulfate)[20], 563 (dimer of oleic acid)[21] and 885 (PI 38:4)[21,22] each illustrated regions of increased ion intensities corresponding to specific histologically evaluated regions of interest (see Figure 1A parts 1-4). Exact mass measurements and fragmentation studies were performed on these ions, as well as others, to support their tentative identifications (Figures S1, S2, Table S1, supplementary information). Sample 36 histological evaluation also identified a region of normal epithelial tissue as illustrated in the top right portion of Figure 1B.

To evaluate the accuracy of DESI-MS in differentiating oral tongue SCC from adjacent normal epithelium, PC1 vs. PC2 Score plots were then created to observe the separation of normal epithelial and SCC average mass spectra within the samples (Figure 2).

Loading plots (Figure 2B and Supplemental Figures 3B, 3D, and 3F) indicate which *m/z* values contribute to separation of samples in score plots (Figure 2A and Supplemental Figures 3A, 3C, and 3E). The loading plots seem to suggest that *m/z* 281, 563, and 885 each aid to separate SCC from normal while *m/z* 465 and 788 each contributes to distinguish normal from SCC. Supplemental Figure 1F suggests that *m/z* 794 (PC 36:1 + Cl⁻)[23] / (PE 40:4)[21] and 885 (PI 38:4)[21,22] each contribute to the variation in normal epithelial and SCC average mass spectra.

Space charged *m/z* 281, 465, and 788 peaks are evident in broadened loading values. In response, we chose to evaluate the maximum value for each *m/z*. As illustrated in Figure 3, One Way ANOVA of TIC normalized normal epithelial and squamous cell carcinoma values at maximum value *m/z* 281 has a p value of $1.15*10^{-3}$, maximum *m/z* 563 has a p value of $4.6*10^{-3}$ and maximum value *m/z* 465 has a p value of $8.55*10^{-5}$. Power analysis provided values of 0.95, 0.99, and 0.98 for *m/z* 281, 563, and 465, respectively. The fold-change in the mean values, normal epithelial tissue relative to SCC, were 0.53, 0.27, and 9.37 for *m/z* 281, 563, and 465, respectively. These univariate results show that *m/z* 281, 563, and 465 may be strong discriminators between normal epithelial tissue and oral SCC and they warrant future study.

PCA-LDA was performed on the DESI-MS imaging averaged mass spectra to distinguish between normal epithelial and SCC. LDA is useful to build a classification model capable of predicting the disease state of unknown samples.[10] The DESI data comprised normal epithelial tissue including eight measurements (n=8) and squamous cell carcinoma including thirteen measurements (n=13). The range of two and four PCs were used for classification. Several combinations of PCs and deletion groups did not significantly increase the highest prediction rates for both classes (normal epithelial and SCC). The range of PCs confirms that the results obtained are relatively stable. The average cross-validation (CV) prediction rate was 95%. The number of false results per class obtained through cross-validation can be seen in Table 1. In Table 1 histopathology is the control testing method and PCA-LDA of DESI-MS imaging is the testing method under evaluation. Multivariate Statistical Analysis, specifically Principal Component Analysis-Linear Discriminant Analysis (PCA-LDA), is used as the method or test applied to DESI-MS imaging data.

PCA-LDA was performed on each selected pixel in each of the sixty-two UC Davis samples. Histopathology regions of normal epithelial tissue and squamous cell carcinoma were correlated with *m/z* 200-1,000 and *m/z* 700-900 pixel predictions. For example, histological evaluation of Sample 43 identified a region of normal epithelial tissue and a region of squamous cell carcinoma as illustrated in Fig. 4C.

As illustrated in Figure 5, both pixel prediction models (*m/z* 200-1,000 and *m/z* 700-900) correctly predict normal epithelial for UC Davis Sample 30 with only a few pixels in the *m/z* 200-1,000 pixel model predicted as squamous cell carcinoma. Both pixel prediction models predicted skeletal muscle as squamous cell carcinoma because the prediction models have only one of two classes to choose from either normal epithelial or squamous cell carcinoma. We expect that creation of additional classes for dysplastic epithelium, skeletal and smooth muscle, fat, connective tissue, and blood vessels will increase prediction rates for heterogeneous tissue samples.

Forty-four of the acquired sixty-two UC Davis samples included either normal epithelial and/or squamous cell carcinoma. Two samples (UC Davis Samples 24 and 43) included both normal epithelial tissue and squamous cell carcinoma providing a total population of 46. All histopathology regions of normal epithelial and squamous cell carcinoma were compared to both pixel validation models. The *m/z* 700-900 pixel prediction model had 91% accuracy while the *m/z* 200-1000 pixel prediction model was 76% accurate (data not shown). In Table 2 histopathology is the control testing method and *m/z* 700-900 pixel validation method by PCA-LDA is the testing method under evaluation.

Discussion

In the current work, DESI-MS imaging of frozen tissue sections detected differentiating MS lipid profiles of SCC and normal samples. We note that this is a first exploration of the differential chemical signals between normal and SCC using DESI-MS imaging to oral cancer. PCA-LDA provided 95% accuracy. One Way ANOVA of TIC normalized normal epithelial and squamous cell carcinoma values at maximum value *m/z* 281, maximum *m/z* 563 and maximum value *m/z* 465 each had p values less than 0.01. The current work benefits by interrogating all tissue biopsies by DESI-MS imaging and histopathological evaluation. The inclusion of comparison of DESI-MS imaging results against pathological evaluation provides confidence in sensitivity and specificity values. The objective data can be collected into databases, sorted by clinical outcome data, and further analyzed to provide more accurate profiles of histopathological tissue types and disease states. The collected data can then be used to minimize patient variability and provide better classification algorithms and improved confidence intervals.

DESI-MS can measure differences in lipid profiles based on lipid based biochemical changes such as fatty acid synthase (FAS). Lipid based biochemical changes, such as overexpression of FAS, have been reported for oral SCC.[24] The DESI-MS imaging data supports this previous finding. FAS plays a central role in the endogenous synthesis of long chain fatty acids from acetyl-CoA and malonyl-CoA.[24] Several human cancers, including oral SCC, constitutively express high levels of FAS and consequently have an active endogenous fatty acid biosynthesis. FAS is higher in oral SCC than in the adjacent morphologically normal epithelium.[24]

In our experiments, the greater relative intensity of cholesterol sulfate (CS) by DESI-MS imaging in normal regions of interest compared to SCC regions of interest (see Figures 1A and 1B part 3) supports previous literature findings. A prior independent study[25] examined abundance of cholesterol sulfate in normal epidermal and squamous cell carcinoma cells in the process of squamous differentiation from growth phase to confluence. At confluence, normal epidermal cells undergo terminal differentiation and expression of the differentiated phenotype except for a variant with defects in terminal differentiation control. Normal epidermal cells except for the variant experienced a greater than 30 times accumulation of cholesterol sulfate. The variant and squamous cell carcinoma each exhibited less than 3 times accumulation of cholesterol sulfate. Their hypothesis is that increased cholesterol sulfate is a marker for squamous differentiation.

Oral cancer, like most carcinomas, arises through progressive histopathologic changes. As a result, these ambient ionization methods should detect differences in lipid-based morphological changes. Current visible and light based detection systems are not as objective and provided less molecular information than MS-based methods. As previously mentioned, DESI-MS imaging lipid profiles in tissue can be correlated with histopathological characterization.[14] DESI-MS analysis of oral cancer has now been shown to do the same. We

also expect DESI-MS to provide diagnostic information before normal epithelium transitions completely from normal to SCC. In other words, we anticipate that mucosa with significant precancerous changes (severe dysplasia or carcinoma in situ) would have DESI-MS profiles different from adjacent normal epithelium.

We acknowledge that the data analyzed so far has been based on analysis of known histopathology results and averaging of selected regions of interest. Furthermore, tissue heterogeneity so far limits the approach to not mixed regions of histopathology, and our study has a relatively limited number of data points. We hypothesize that, similar to other DESI-MS based oncology studies, larger data sets and more detailed pathological evaluation will improve classification of SCC and normal samples as well as differentiate adjacent normal epithelial samples. The ultimate test being the application of this methodology on samples of unknown histopathology and comparison of DESI-MS prediction with that of traditional histopathology.

For the pixel prediction models, histopathology regions of normal epithelial and squamous cell carcinoma were compared to *m/z* 200-1,000 and *m/z* 700-900 pixel validation models. As illustrated in Figures 4 – 5, both pixel prediction models (*m/z* 200-1,000 and *m/z* 700-900) typically provided a correct prediction of squamous cell carcinoma and normal epithelial tissue. The prediction models, however, are limited to two classifications (SCC or normal epithelium) and this limits the successful application of the technique. Creation of additional classes (such as skeletal muscle, fat, fibrous tissue, etc.) should increase prediction rates for heterogeneous tissue samples.

Overall *m/z* 700-900 pixel validation model by PCA-LDA provided 91% accuracy. More work is needed to confirm this method and classification of all typical tissue types and pixel predictions with additional classifications should increase confidence in the validation model. Analysis of different *m/z* ranges, such as *m/z* 200-400 (which includes oleic acid at *m/z* 281) and 400-600 (which includes cholesterol sulfate at *m/z* 465 and oleic acid dimer at *m/z* 563) may provide high accuracy rates. Additional *m/z* ranges may confirm existing validation models.

On top of these challenges, some additional issues persist. As illustrated in Figure 6, both pixel prediction models (*m/z* 200-1,000 and *m/z* 700-900) fail to correctly predict regions of normal epithelial tissue adjacent to squamous cell carcinoma (UC Davis Sample 24). Similarly, as illustrated in supplemental Figure 5, both pixel prediction models (*m/z* 200-1,000 and *m/z* 700-900) incorrectly predict normal epithelial in one sample (UC Davis Sample 9). Analysis of the average mass spectra can be used to explain the normal epithelial tissue prediction but not why the prediction was incorrect. The intermittent inability to correctly predict normal epithelial tissue adjacent to squamous cell carcinoma in oral samples has been a consistent limitation of this methodology. This is likely due to the fact that adjacent tissue, while normal in appearance histologically, actually demonstrates premalignant changes on a molecular level, thus signaling that it may be at risk for future carcinogenesis. Moreover, squamous cell carcinoma can have differing degrees of differentiation that may lead to different cellular lipid profiles, possibly making it more or less easily discerned from adjacent normal epithelium. Creation of an additional class for normal epithelial tissue adjacent to squamous cell carcinoma, as well as different norms for the spectrum of premalignant change (severe dysplasia and carcinoma *in situ*) and tumor differentiation, should increase prediction accuracy through a larger dataset of these different categories. This may be an area of future study.

Conclusion

This work suggests that DESI-MS may be useful in the diagnosis of oral tongue squamous cell carcinoma. The time for DESI-MS imaging (approx. 20 min. per sample) is not compatible with point-of-care use. Transitioning to sparse DESI-MS analysis of tissue smears, as shown in previous work with brain cancer[16] and independent studies with breast cancer,[26] would increase sample throughput and allow intraoperative use. Also, the direct MS analysis of oral brush biopsies,[27] a routine method in oral cancer screening, using touch spray mass spectrometry[28] would be an alternative to DESI that would provide similar diagnostic information[29] on a short timescale. The time to collect a single mass spectrum is on the order of milliseconds and statistical classification could be reduced to about a minute per sample. Further work is needed to validate this method prospectively, and to develop techniques for evaluating the technology for use in intraoperative margin assessment as well as in oral cavity cancer screening.

5   Conflict of Interest

No potential conflicts of interest were disclosed.

References

1. Bonello, L., *et al.* Squamous cell carcinoma of the oral cavity and oropharynx: what does the apparent diffusion coefficient tell us about its histology? *Acta Radiol* (2015).
2. Dey, K.K., *et al.* Identification of RAB2A and PRDX1 as the potential biomarkers for oral squamous cell carcinoma using mass spectrometry-based comparative proteomic approach. *Tumour Biol* (2015).
3. Shah, F.D., *et al.* A review on salivary genomics and proteomics biomarkers in oral cancer. *Indian J Clin Biochem* **26**, 326-334 (2011).
4. Lingen, M.W., Kalmar, J.R., Karrison, T. & Speight, P.M. Critical evaluation of diagnostic aids for the detection of oral cancer. *Oral Oncol* **44**, 10-22 (2008).
5. Lippman, S.M. & Hong, W.K. Second malignant tumors in head and neck squamous cell carcinoma: the overshadowing threat for patients with early-stage disease. *Int J Radiat Oncol Biol Phys* **17**, 691-694 (1989).
6. Fedele, S. Diagnostic aids in the screening of oral cancer. *Head & neck oncology* **1**, 5 (2009).
7. Ramadas, K., *et al.* Interim results from a cluster randomized controlled oral cancer screening trial in Kerala, India. *Oral Oncol* **39**, 580-588 (2003).
8. Barrellier, P., Babin, E., Louis, M.Y. & Meunier-Guttin, A. [The use of toluidine blue in the diagnosis of neoplastic lesions of the oral cavity]. *Rev Stomatol Chir Maxillofac* **94**, 51-54 (1993).
9. Lane, P.M., *et al.* Simple device for the direct visualization of oral-cavity tissue fluorescence. *J Biomed Opt* **11**, 024006 (2006).
10. Kerian, K.S., *et al.* Differentiation of prostate cancer from normal tissue in radical prostatectomy specimens by desorption electrospray ionization and touch spray ionization mass spectrometry. *Analyst* **140**, 1090-1098 (2015).
11. Pereira, M.C., Oliveira, D.T., Landman, G. & Kowalski, L.P. Histologic subtypes of oral squamous cell carcinoma: prognostic relevance. *J Can Dent Assoc* **73**, 339-344 (2007).
12. Nagaraj, N.S. Evolving 'omics' technologies for diagnostics of head and neck cancer. *Brief Funct Genomic Proteomic* **8**, 49-59 (2009).
13. Tung, C.L., *et al.* Proteomics-based identification of plasma biomarkers in oral squamous cell carcinoma. *J Pharm Biomed Anal* **75**, 7-17 (2013).
14. Eberlin, L.S., Ferreira, C.R., Dill, A.L., Ifa, D.R. & Cooks, R.G. Desorption electrospray ionization mass spectrometry for lipid characterization and biological tissue imaging. *Biochim Biophys Acta* **1811**, 946-960 (2011).
15. Eberlin, L.S., *et al.* Ambient mass spectrometry for the intraoperative molecular diagnosis of human brain tumors. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 1611-1616 (2013).
16. Pirro, V., *et al.* Intraoperative assessment of tumor margins during glioma resection by desorption electrospray ionization-mass spectrometry. *Proceedings of the National Academy of Sciences of the United States of America* **114**, 6700-6705 (2017).
17. Dill, A.L., *et al.* Multivariate statistical identification of human bladder carcinomas using ambient ionization imaging mass spectrometry. *Chemistry* **17**, 2897-2902 (2011).
18. Ifa, D.R., Wiseman, J.M., Song, Q.Y. & Cooks, R.G. Development of capabilities for imaging mass spectrometry under ambient conditions with desorption electrospray ionization (DESI). *Int J Mass Spectrom* **259**, 8-15 (2007).
19. Wiseman, J.M., Ifa, D.R., Venter, A. & Cooks, R.G. Ambient molecular imaging by desorption electrospray ionization mass spectrometry. *Nat. Protocols* **3**, 517-524 (2008).
20. Pirro, V., *et al.* Lipid characterization of individual porcine oocytes by dual mode DESI-MS and data fusion. *Anal Chim Acta* **848**, 51-60 (2014).
21. Girod, M., Shi, Y., Cheng, J.X. & Cooks, R.G. Desorption electrospray ionization imaging mass spectrometry of lipids in rat spinal cord. *J Am Soc Mass Spectrom* **21**, 1177-1189 (2010).
22. Eberlin, L.S., *et al.* Discrimination of human astrocytoma subtypes by lipid analysis using desorption electrospray ionization imaging mass spectrometry. *Angew Chem Int Ed Engl* **49**, 5953-5956 (2010).

23.   Zhang, X. & Reid, G.E. Multistage tandem mass spectrometry of anionic phosphatidylcholine lipid adducts reveals novel dissociation pathways. *Int J Mass Spectrom* **252**, 242-255 (2006).

24.   Silva, S.D.*, et al.* Fatty acid synthase expression in squamous cell carcinoma of the tongue: clinicopathological findings. *Oral Dis* **14**, 376-382 (2008).

25.   Jetten, A.M., George, M.A., Nervi, C., Boone, L.R. & Rearick, J.I. Increased cholesterol sulfate and cholesterol sulfotransferase activity in relation to the multi-step process of differentiation in human epidermal keratinocytes. *The Journal of investigative dermatology* **92**, 203-209 (1989).

26.   Woolman, M.*, et al.* An Assessment of the Utility of Tissue Smears in Rapid Cancer Profiling with Desorption Electrospray Ionization Mass Spectrometry (DESI-MS). *J Am Soc Mass Spectrom* **28**, 145-153 (2016).

27.   Babshet, M., Nandimath, K., Pervatikar, S. & Naikmasur, V. Efficacy of oral brush cytology in the evaluation of the oral premalignant and malignant lesions. in *J Cytol*, Vol. 28 165-172 (2011).

28.   Kerian, K.S., Jarmusch, A.K. & Cooks, R.G. Touch spray mass spectrometry for in situ analysis of complex samples. *Analyst* **139**, 2714-2720 (2014).

29.   Pirro, V.*, et al.* Analysis of human gliomas by swab touch spray-mass spectrometry: applications to intraoperative assessment of surgical margins and presence of oncometabolites. *Analyst* (2017).
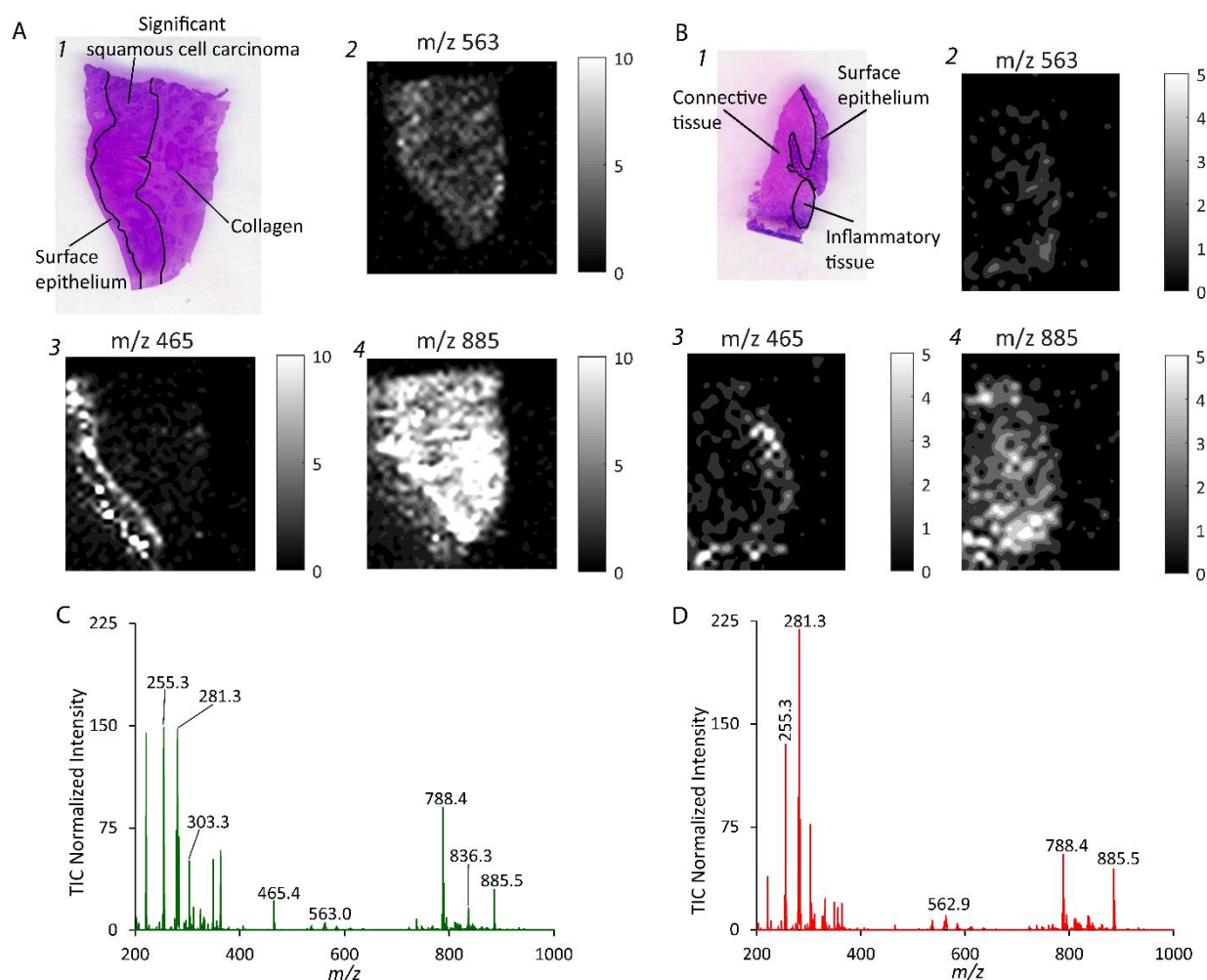
Figure 1. Histopathological evaluation (performed by DJS) and DESI-MS data for select samples. **A&B)** 1 corresponds to H&E stained image with pathological evaluation, 2-4 are ion images of *m/z* 563, 465, 885, on a grayscale with corresponding scale bar. **A.** Sample 35. **B.** Sample 36. Scanned image of H&E stained Sample 36 with histopathology determined annotations delineating regions of interest. **B.** Scanned image of H&E stained Sample 35 with histopathology determined annotations delineating regions of interest. **C.** Average mass spectrum, normalized to the TIC and scaled by 100, of selected pixels of interest for Sample 35 normal epithelial tissue. **D.** TIC normalized and scaled by 100 average mass spectrum of selected pixels of interest for Sample 35 (squamous cell carcinoma with collagen).
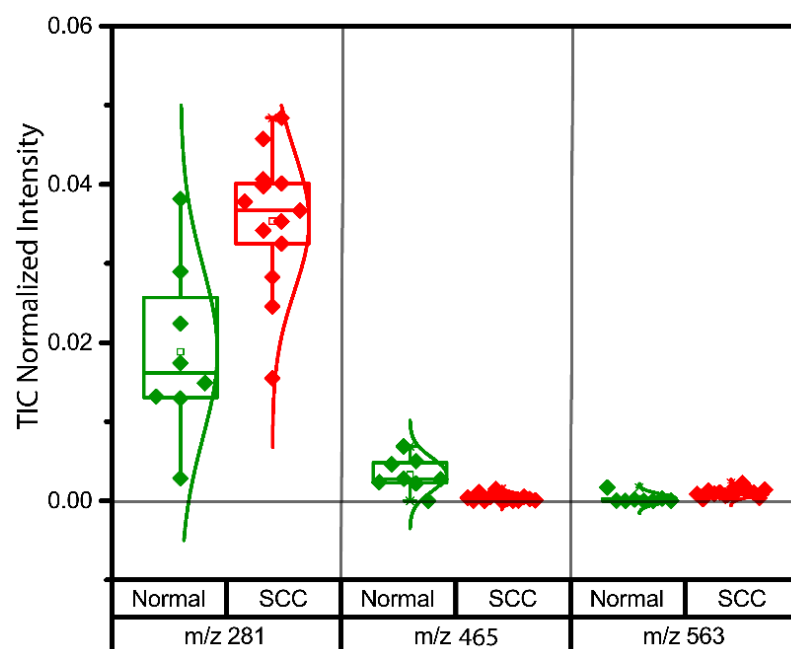
Figure 2. A. PCA score plot (PC1 vs. PC2) displays normal epithelial from SCC average mass spectra using DESI-MS imaging negative mode *m/z* range 400 - 600. SCC average mass spectra are red triangles and normal epithelial average mass spectra are green circles. **B.** PCA loading plot (PC1 vs. PC2) showing differentiation based on *m/z* 465 (cholesterol sulfate)[19] (positive PC1 coefficients and positive and negative PC2 coefficients) and to a lesser extent *m/z* 563 (dimer of oleic acid)[20] (negative PC1 and PC2 coefficients).

Figure 3. Comparison of TIC Normalized (without scaling) maximum values of *m/z* 281, 465 and 563 between normal epithelial (green) and SCC (red) from the DESI-MS imaging data.
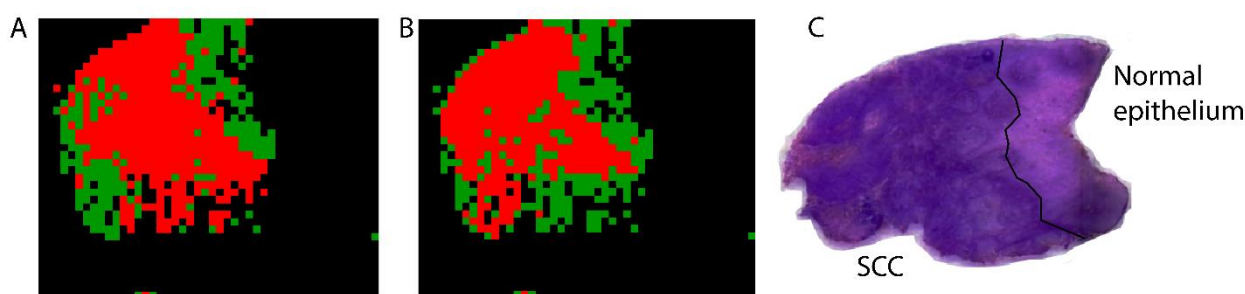
Figure 4. A. Pixel prediction for *m/z* 200-1,000 for Sample 43. Red predicting squamous cell carcinoma and green predicting non-SCC. **B.** Pixel prediction for *m/z* 700-900 for Sample 43. **C.** Histopathology determination for Sample 43.
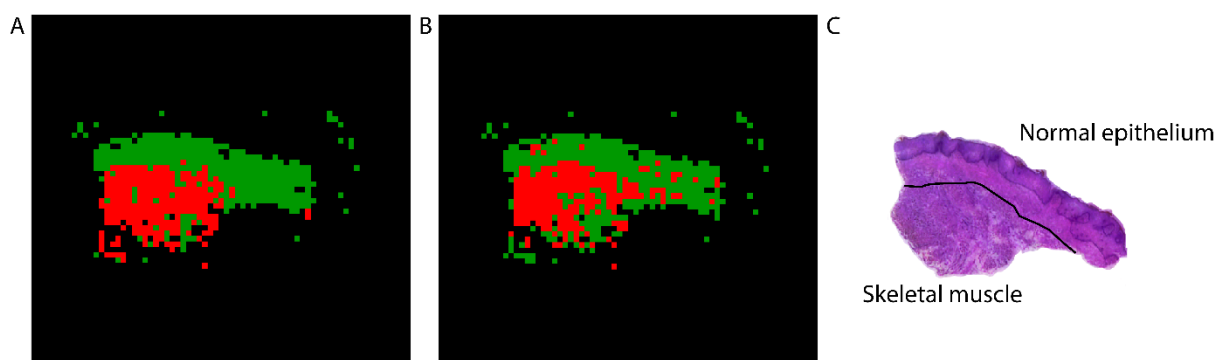

Figure 5. A. Pixel prediction for *m/z* 200-1,000 for UC Davis Sample 30. Red predicting squamous cell carcinoma and green predicting normal epithelial tissue. **B.** Pixel prediction for *m/z* 700-900 for UC Davis Sample 30. **C.** Histopathology determination of normal epithelium and skeletal muscle for UC Davis Sample 30.
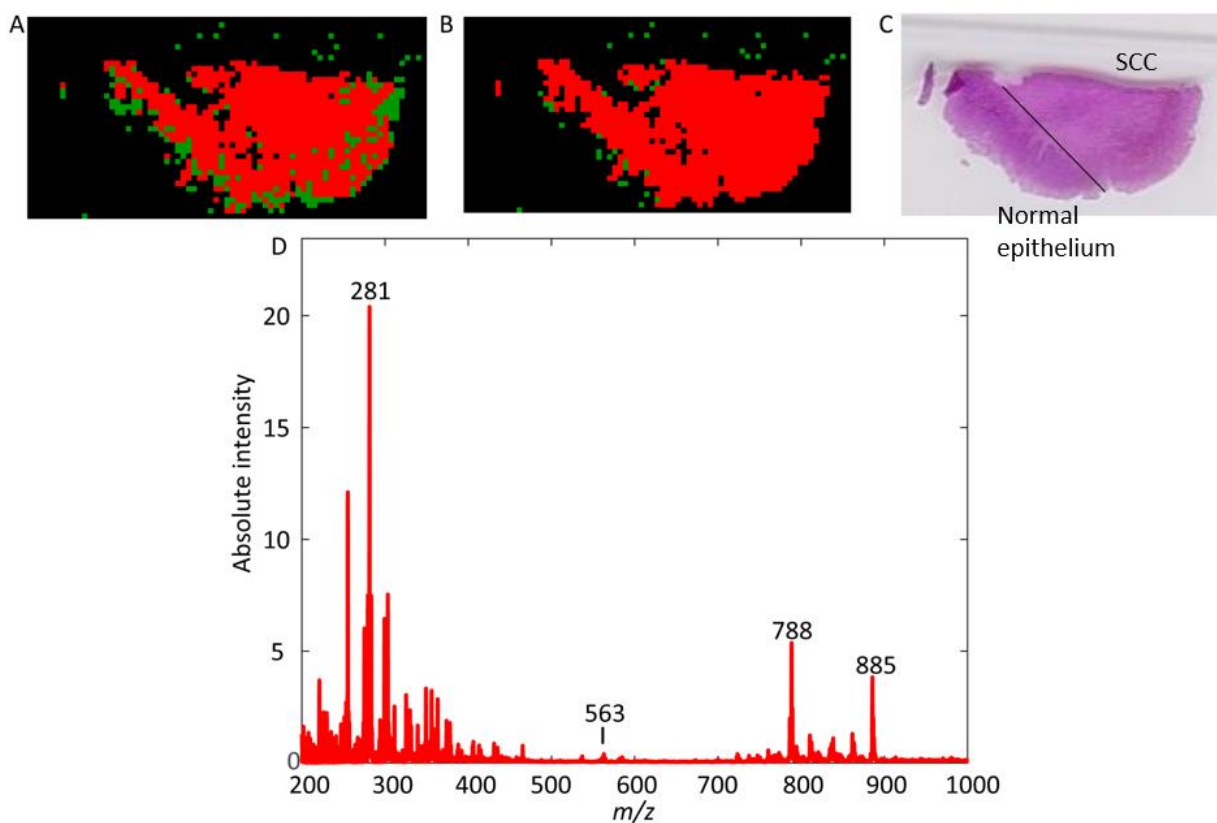
Figure 6. A. Pixel prediction for *m/z* 200-1,000 for UC Davis Sample 24. Red predicting squamous cell carcinoma and green predicting non-SCC. **B.** Pixel prediction for *m/z* 700-900 for UC Davis Sample 24. **C.** Histopathology determination of squamous cell carcinoma for UC Davis Sample 24. **D.** Average mass spectrum of selected pixels for UC Davis Sample 24.

Table 1: Cross Validation Confusion Matrix for DESI-MS imaging Prediction Rates:

Total Population (21 samples)		Pathology		SCC Prevalence (62%, 13/21)
		SCC	Normal	
DESI-MS imaging PCA-LDA	SCC	13	1	Positive Predictive Value 92.8%
	Normal	0	7	Negative Predictive Value 100%
	Accuracy (95%)	True Positive Rate 100%	True Negative Rate 87.5%	

Pixel by Pixel Prediction Models of Oral Cancer Samples

Table 2: Confusion Matrix for m/z 700-900 DESI-MS imaging Prediction Rates:

Total Population (46 samples)		Pathology		SCC Prevalence (63%, 29/46)
		SCC	Normal	
m/z 700-900 Pixel by Pixel PCA-LDA	SCC	27	2	Positive Predictive Value 93%
	Normal	2	15	Negative Predictive Value 88%
	Accuracy (91%)	True Positive Rate 93%	True Negative Rate 88%	