
Article:
McRae, M., Modak, S., Simmons, G. et al. (10 more authors) (2020) Point-of-care oral cytology tool for the screening and assessment of potentially malignant oral lesions. Cancer Cytopathology. ISSN 1934-662X

https://doi.org/10.1002/cncy.22236

Title

Point-of-Care Oral Cytology Tool for the Screening and Assessment of Potentially Malignant Oral Lesions

Running Title

Point-of-Care Oral Cytology Tool

Authors

1. Michael P. McRae, Ph.D.
   Department of Biomaterials, Bioengineering Institute, New York University, New York, NY, USA
   michael.mcrae@nyu.edu
2. Sayli S. Modak, M.S.
   Department of Biomaterials, Bioengineering Institute, New York University, New York, NY, USA
   sayli.modak@nyu.edu
3. Glennon W. Simmons, B.S.
   Department of Biomaterials, Bioengineering Institute, New York University, New York, NY, USA
   glennon.simmons@nyu.edu
4. Denise A. Trochesset, D.D.S.
   Department of Oral and Maxillofacial Pathology, Radiology & Medicine, New York University College of Dentistry, New York, NY, USA
   dat5@nyu.edu
5. A. Ross Kerr, D.D.S., M.S.D.
   Department of Oral and Maxillofacial Pathology, Radiology & Medicine, New York University College of Dentistry, New York, NY, USA
   ark3@nyu.edu
6. Martin H. Thornhill, M.B.B.S., B.D.S., Ph.D.
   Department of Oral & Maxillofacial Medicine, Surgery and Pathology, School of Clinical Dentistry, University of Sheffield, Sheffield, UK
   M.Thornhill@sheffield.ac.uk
7. Spencer W. Redding, D.D.S., M.Ed.
   Department of Comprehensive Dentistry and Mays Cancer Center, The University of Texas Health Science Center at San Antonio, San Antonio, TX, USA
   redding@uthscsa.edu
8. Nadarajah Vigneswaran, B.D.S., Dr.Med.Dent., D.M.D.

Department of Diagnostic and Biomedical Sciences, The University of Texas Health Science Center at Houston, Houston, TX, USA
Nadarajah.Vigneswaran@uth.tmc.edu

9. Stella K. Kang, M.D., M.S.

Departments of Radiology, Population Health New York University School of Medicine, New York, NY, USA

stella.kang@nyulangone.org

10. Nicolaos J. Christodoulides, Ph.D.

Department of Biomaterials, Bioengineering Institute, New York University, New York, NY, USA
nicolaoschristo19@gmail.com

11. Craig Murdoch, Ph.D.

Department of Oral & Maxillofacial Medicine, Surgery and Pathology, School of Clinical Dentistry, University of Sheffield, Sheffield, UK

c.murdoch@sheffield.ac.uk

12. Steven J. Dietl, M.Eng.

SensoDx, LLC, Victor, NY, USA
sdietl@sensodx.com

13. Roger Markham, Ph.D.
SensoDx, LLC, Victor, NY, USA
rmarkham@sensodx.com

14. John T. McDevitt, Ph.D.
Chair, Department of Biomaterials, Bioengineering Institute, New York University, New York, NY, USA
mcdevitt@nyu.edu

Corresponding Author: John T. McDevitt, Ph.D., Chair, Department of Biomaterials, Bioengineering Institute, New York University, 433 First Avenue, Room 820, New York, NY 10010-4086, USA. Email: mcdevitt@nyu.edu. Phone: 212-998-9204.

Funding Support

Funding for this work was provided by the National Institutes of Health (NIH) through the National Institute of Dental and Craniofacial Research (Award Number 1RC2DE020785-01, 5RC2DE020785-02, 3RC2DE020785-02S1, 3RC2DE020785-02S2, 4R44DE025798-02, R01DE024392). The content of this paper is solely the responsibility of the authors and does not necessarily represent or reflect the official

views of the NIH or the US government. Segments of this work are supported by Renaissance Health Service Corporation and Delta Dental of Michigan.

Conflict of Interest Disclosures

Principal Investigator, John T. McDevitt, has an equity interest in SensoDx, LLC. He also serves on the Scientific Advisory Board of SensoDx. Michael P. McRae has served as a consultant for SensoDx.

Author Contributions

Michael P. McRae: Conceptualization, investigation, data curation, formal analysis, methodology, software, validation, visualization, writing - original draft, and writing - review and editing. **Sayli S. Modak:** Investigation, validation, and writing - review and editing. **Glennon W. Simmons:** Project administration, investigation, methodology, data collection, and writing - review and editing. **Denise A. Trochesset:** Conceptualization, writing - review and editing. **A. Ross Kerr:** Conceptualization and writing - review and editing. **Martin H. Thornhill:** Conceptualization and writing - review and editing. **Spencer W. Redding:** Conceptualization and writing - review and editing. **Nadarajah Vigneswaran:** Conceptualization and writing - review and editing. **Stella K. Kang:** Conceptualization and writing - review and editing. **Nicolaos J. Christodoulides:** Conceptualization, writing - original draft, and writing - review and editing. **Craig Murdoch:** Investigation and writing - review and editing. **Steven J. Dietl:** Methodology and writing - review and editing. **Roger Markham:** Methodology and writing - review and editing. **John T. McDevitt:** Conceptualization, funding acquisition, project administration, resources, supervision, writing - original draft, and writing - review and editing.

Acknowledgements

The authors thank the University of Texas Health Science Center at San Antonio (UTHSCSA) (Stephanie Rowan, Chih-Ko Yeh, Stan McGuff, Frank Miller), University of Texas Health Science Center at Houston (UTHSCH) (Jerry Bouquot, Nagi Demian, Etan Weinstock, Nancy Bass), New York University / Bluestone Center for Clinical Research (Joan Phelan, Patricia Corby, Ismael Khouly), Sheffield Teaching Hospitals NHS Foundation Trust and the University of Sheffield (Paul Speight, Christine Freeman, Anne Hegarty, Katy D'Apice) for assistance in obtaining clinical samples. The authors also thank Rho, Inc. (Chapel Hill,

North Carolina) (Julie Vick) for assisting with patient data management and (Robert James) for statistical and data analysis support. Finally, the authors thank Timothy J. Abram for early contributions to the project, including assay development and database organization.

Precis

A point-of-care oral cytology tool was developed for non-invasive detection and monitoring of potentially malignant oral lesions. Distributions of cell phenotypes identified by machine learning and a cytology-on-a-chip approach provide useful information in the assessment of oral lesions with improved interpretability, calibration, and generalizability relative to conventional methods.

Abstract

BACKGROUND: Effective detection and monitoring of potentially malignant oral lesions (PMOL) are critical to identifying early stage cancer and improving outcomes. In this study, the authors describe cytopathology tools including machine learning algorithms, clinical algorithms, and test reports developed to assist pathologists and clinicians with PMOL evaluation. **METHODS:** Data were acquired from a multi-site clinical validation study of 999 subjects with PMOLs and oral squamous cell carcinoma (OSCC) using a cytology-on-a-chip approach. A machine learning model was trained to recognize and quantify the distributions of four cell phenotypes. A least absolute shrinkage and selection operator (lasso) logistic regression model was trained to distinguish PMOLs and cancer across a spectrum of histopathologic diagnoses ranging from benign, to increasing grades of oral epithelial dysplasia (OED), to OSCC using demographics, lesion characteristics, and cell phenotypes. Cytopathology software was developed to assist pathologists in reviewing brush cytology test results, including high-content cell analyses, data visualization tools, and results reporting. **RESULTS:** Cell phenotypes were accurately determined through an automated cytological assay and machine learning approach (99.3% accuracy). Significant differences in cell phenotype distributions across diagnostic categories were found in three phenotypes (Type 1 'mature squamous', Type 2 'small round', and Type 3 'leukocytes'). The clinical algorithms resulted in acceptable performance characteristics (AUC = 0.81 for benign vs. mild dysplasia and 0.95 for benign vs. malignancy). **CONCLUSION:** These new cytopathology tools represent a practical solution for rapid PMOL assessment

with the potential to facilitate screening and longitudinal monitoring in primary, secondary, and tertiary clinical care settings.

KEY WORDS: squamous cell carcinoma; oral epithelial dysplasia; point-of-care testing; single-cell analysis; artificial intelligence; cytology; biomarkers.

Text Pages: 23

Tables: 0

Figures: 7

Supporting Files for Publication: Suppl. Methods, Suppl. Figures (6 total), Suppl. Video Content

Introduction

Cancers of the lip, oral cavity, and pharyngeal subsites are estimated to affect over 500,000 people globally each year.[1] The National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) program estimates 53,000 new cases and 10,860 deaths attributed to oral and pharyngeal cancer (OPC) in 2019 in the US alone, of which approximately 50% involve oral cavity subsites. Collectively, OPCs represent approximately 3% of all cancers.[2] Approximately two-thirds of OPCs are diagnosed at Stage III or IV when the 5-year survival rate is just 45% and 32%, respectively.[3] For the remaining third of OPCs detected at early stages,[4] survival increases to 84%.[2] Despite steady improvements in overall survival rates for OPC over the last four decades, identifying OPCs at an early stage remains a challenge for oral health care providers.[5] The current diagnostic paradigm of procuring a biopsy is based on remote lab services which can take days/weeks to provide results, and this further prolongs anxiety for patients. A point-of-care (POC) solution could provide immediate feedback within the same visit. Thus, there is a strong need for technology-driven solutions that can precisely and rapidly diagnose the entire spectrum of oral epithelial dysplasia (OED) and oral squamous cell carcinoma (OSCC) using minimally invasive sampling at the POC.

A successful diagnostic adjunctive test for primary care settings should be able to discriminate potentially malignant oral lesions (PMOLs) that are at "risk" (i.e., malignant lesions or those with an elevated risk for undergoing malignant transformation) from more common benign lesions with no malignant potential, thus improving the referral efficiency to secondary or tertiary care (e.g., reducing over-referral of patients with benign lesions and improving the early identification and prompt referral of malignant or high-grade dysplastic PMOLs for oncologic care). Numerous adjunctive tests are available to assist in the diagnosis of PMOLs. In a meta-analysis of oral cancer adjuncts, vital staining and visualization adjuncts (e.g., autofluorescence and tissue reflectance) demonstrated insufficient accuracy to be recommended for use as lesion triage tools by general dentists.[6] Cytology, however, has demonstrated greater sensitivity and specificity relative to the other adjuncts, suggesting its potential as a surrogate for gold-standard histopathology. This evidence to support the accuracy of cytology is largely based on accuracy studies performed in secondary and tertiary care settings. Although cytology is unable to replace histopathologic diagnosis based on tissue architecture, this relatively inexpensive, easy to perform, and minimally-invasive method may be useful for triaging lesions in any setting: primary care settings such as a dental office, low-

resource/remote settings, and secondary/tertiary settings. Incisional biopsy followed by histopathologic examination represents the current standard of care for diagnosing PMOLs. However, incisional biopsy of PMOLs, particularly in those that are large non-homogeneous leukoplakias, leads to underestimation of the severity of OED up to 30% of the time because the biopsy sample (typically 5 mm in diameter) may not be representative of the variable pathology across the field of the entire PMOL.[7] Brush cytology could enable a wider sampling of PMOLs that encompass larger areas or are multifocal with the potential to reduce sampling errors encountered with incisional biopsies.

Previously, we have demonstrated the conceptual basis and the efficacy of chip-based cell capture, multispectral fluorescence measurements, and single-cell analysis approaches yielding high content diagnostic information related to oral lesions.[8-10] This compact and integrated lesion diagnostic adjunct approach has been studied previously through a multi-site clinical validation effort that has led to the development of one of the largest oral cytology databases ever assembled for PMOLs.[11,12] These efforts included the development of an “enhanced gold standard” adjudication process[12] that was used to correlate brush cytology measurements with six levels of histopathological diagnosis, ranging from benign, to OED, to OSCC. The same approach showed strong promise for OSCC surveillance in Fanconi Anemia patients[13] and for the development of a cytology based numerical risk index for cancer progression.[14] Overall, these past efforts have revealed that microfluidic-based cell capture systems with integrated imaging and embedded diagnostic algorithms can yield diagnostic accuracies that rival and exceed the capabilities of previously developed adjunct devices. These tools were developed previously to serve as adjunctive aids capable of distinguishing between high risk and low risk oral lesions with the goal of improving the pipeline of referrals from primary care settings to secondary and tertiary treatment centers. Thus, these models were intended for assisting primary care providers in making binary referral decisions and considered hundreds of complicated image-based cytomorphometric features with minimal clinical interpretability (i.e., “black box”).

This manuscript targets the development of a Point of Care Oral Cytology Tool (POCOCT), the first precision oncology technology capable of high content cell analysis for near patient testing. The POCOCT platform comprises a minimally invasive brush cytology test kit, disposable assay cartridge, instrument, clinical algorithms, and cloud-based software services that automate the quantification and analysis of

cellular and molecular signatures of dysplasia with results available in a matter of minutes as compared to days for traditional labor intensive lab-based pathology methods. This paper features the development of new diagnostic models using the same database described above with the goal of greatly simplifying the diagnostic algorithms and their interpretation through the classification and quantification of cellular phenotypes, resulting in more informative and transparent models for cytopathologists. Likewise, this work explores the utility of cell phenotype identification through machine learning, their implementation in diagnostic models with interpretable predictors and responses, and the practical application of these software tools in a cytopathology service.

Materials and Methods

Oral Cytology Data

Data used in this study originated from the 999-patient multisite prospective non-interventional study evaluating the cytology-on-a-chip system for the measurement of cytological parameters on brush cytology samples to assist in the diagnosis of PMOL.[11,12] Briefly, both histopathological and brush cytological samples for 714 subjects from three patient groups were measured: (1) subjects with PMOL who underwent scalpel biopsy as part of the standard of care for microscopic diagnosis, (2) subjects with recently diagnosed malignant lesions, and (3) healthy volunteers without lesions. Histopathological assessment of scalpel biopsy specimens classified lesions into six categories (benign, mild-, moderate- or severe-dysplasia, carcinoma-in-situ, and OSCC), including healthy controls without lesions. While traditionally the grading of OED has been considered subjective and lacking intra- and inter-observer reproducibility,[15,16] this new study implemented an “enhanced gold standard” adjudication.[12] Here, two adjacent serial histologic sections were independently scored by two pathologists. In the event that the pathologists disagreed, a third independent adjudicating pathologist reviewed both sections. If the adjudicator did not agree with either of the initial two pathologists, a third stage consensus review was conducted to attain a final diagnosis. This “enhanced gold standard” process was able to achieve 100% consensus agreement compared to an initial pre-adjudication 69.9% agreement rate.

Brush cytology specimens were collected and processed using protocols published previously.[11,12] Cytopathological assessment of brush cytology specimens implemented a cytology-on-a-chip approach

which measured morphological and intensity-based cell metrics as well as the expression of six molecular biomarkers (αvβ6, EGFR, CD147, McM2, Geminin, and Ki67), resulting in a total of 13 million cells analyzed with over 150 image-based parameters. The molecular biomarkers were selected based on their capacity to distinguish benign, dysplastic, and malignant oral epithelial cells through prior immunohistochemistry studies.[9,17,18] Specific details on the molecular biomarker selection, patient characteristics, sample collection and processing, cytology assay, and cytological parameters were published previously[11] and are summarized in the **Supplemental Methods**.

Cell Identification Model Training and Validation

A cell phenotype classification model was explored for its ability to discriminate and quantitate the frequency and distributions of four cell phenotypes: Type 1: cells presenting as polygonal in shape with a low nuclear-cytoplasmic ratio (NC ratio) which represent mature squamous epithelial cells; Type 2: cells presenting as small round cells representing immature parabasal cells; Type 3: cells presenting as mononuclear leukocytes; Type 4: cells represented by lone (naked) nuclei without cell membrane and cytoplasm. To recognize these cell types, a machine learning algorithm was trained on 144 cellular/nuclear features from single-cell analyses, including morphological and intensity-based measurements. Prior to model development, principal component analysis (PCA) was performed on the training set. The PCA method is an unsupervised statistical learning technique for exploratory data analysis which improves data visualization by reducing the dimensionality of complex datasets[19] and has been used for phenotypic identification in flow cytometric data.[20] Detailed methods for the training and validation of the cell identification model are provided in the **Supplemental Methods**.

Numerical Index and Diagnostic Models for Assessing PMOL

A numerical index was developed for the purpose of discriminating benign vs. dysplasia/malignant lesions (OED-spectrum model 2|3). Detailed methods for the training and validation of the numerical index and detailed definition of predictors are provided in the **Supplemental Methods**. Briefly, subjects were dichotomized into “case” and “non-case” outcomes according to their lesion determination (non-case for benign lesions and case for [mild, moderate, severe] dysplasia and malignant lesions). Due to relatively few numbers of moderate and severe dysplasia patients (total of 21), these lesion determinations were

combined. Lasso logistic regression was selected for its ability to reduce the number of predictors in high-dimensional datasets to improve prediction performance and generalizability.[21-24] Non-zero lasso logistic regression coefficients were retained for the following predictors: percentage of non-mature squamous cells, percentage of small round cells, percentage of leukocytes, age, sex, smoking pack years, lesion major axis diameter, clinical impression of lichen planus, and lesion color (red, white, or red/white). Diagnostic performance was characterized by area under the curve (AUC), sensitivity, and specificity. Median numerical indices were compared for each diagnostic classification using a two-sided Wilcoxon rank sum test at a significance level of $p = 0.05$. Internal calibration was performed by sorting and grouping the predicted responses (i.e., numerical index) into deciles and measuring the observed proportions of dysplasia/malignant lesions in each decile. The Hosmer-Lemeshow goodness of fit statistic was used to assess the model fit.[21]

Following this same method, diagnostic algorithms for mild vs. moderate dysplasia (OED-spectrum model 3|4), low vs. high risk (4|4), moderate vs. severe dysplasia (4|5), healthy control (no lesion) vs. malignant (0|6), and benign vs. malignant (2|6) were also developed, and AUC, sensitivity, and specificity were reported as mean and 95% confidence interval values for the cross-validated test set.

Cytopathology Software

Measurements of individual cells, such as morphometric appearance and biomarker staining intensity, were recorded using the open-source software CellProfiler.[25] All model development and data analyses were completed with MATLAB R2017b (MathWorks, Natick, MA, USA) software. A graphical user interface for visualizing cytopathology results was developed in MATLAB R2017b. The results summary report tool was developed with Python 3.6.3. Figures of the cytopathology software interface and results summary were compiled from a test on the integrated POCOCT instrument.

Level of Integration

Data originating from our 999-patient NIH Grand Opportunity (GO) study and used in the cell identification and diagnostic models were collected using non-integrated cytology-on-a-chip flow cell prototypes, syringe pumps, research microscope stations, and a collection of commercial and open-source software packages (see **Supplemental Methods** for more details).[11] More recently, we have integrated the cytology-on-a-chip technology into a POC device comprising integrated instrument, microfluidic cartridges

with on-board blister packs, and dedicated software. Likewise, sample processing steps have been significantly reduced. Cell identification and diagnostic models developed on the non-integrated platform were translated to the POC instrument, and software screenshots and results reports presented here were completed with this integrated POC platform.

Results

Cell Identification Model

A cell identification tool to assist in the accurate and precise estimation of histopathological endpoints for the entire spectrum of OED and OSCC was developed. Figure 1 shows the diagnostic categories and rates for oral cancer and dysplasia based on WHO classification[26] found during mass screening,[27] showing 5-year malignant transformations[28] and 5-year cancer recurrence.[29] The literature presents a range of 5-year transformation and recurrence rates, and the ones listed here are representative of those reported previously.[30]

The POCOCT platform (Figure 2) comprises a minimally invasive brush cytology test kit, disposable assay cartridge, instrument, clinical algorithms, and cloud-based software services to automate the quantification and analysis of cellular and molecular signatures of dysplasia and OSCC. The cell identification tool automatically classified four distinct cell phenotypes (Figure 3A). Type 1 'mature squamous' or 'mature keratinocytes' were broad/flat cells, approximately 50-100 µm in diameter, had a low NC ratio, and demonstrated a relatively low cytoplasm staining intensity (Phalloidin-Alexa Fluor® 647). Type 2 'small round' cells were small (12-30 µm in diameter) highly circular cells with high NC ratio and a brightly stained cytoplasm representing immature basaloid keratinocytes. Type 3 'leukocytes' appeared as small, brightly stained pink objects 6-23 µm in diameter representing mononuclear leukocytes. Type 4 'lone nuclei' represented by lone or naked nuclei without a cytoplasm appeared as brightly stained blue objects approximately 5-12 µm in diameter.

The PCA scatter plot of the first two principal components revealed a glimpse of the internal data structure and variance (Figure 3A). Here, populations according to each cell type were clearly observed. Further, over 90% of the variance was explained by the first 20 principal components from a total of 144, with 30% and 14% variance explained in the first and second principal components, respectively. Despite

Types 2 and 3 having similar cytomorphology, the features with the largest association with the first principal component were NC ratio and mean cytoplasm intensity, suggesting that cell size and cellular actin content/distribution play a dominant role in explaining the variance among these cell phenotypes.

The cross-validated *k-nearest neighbors (k*-NN) algorithm resulted in overall accuracy of 96.9% and accuracy of 100%, 90.1%, 96.0%, and 99.0% for Types 1 (mature), 2 (small), 3 (leukocytes), and 4 (lone nuclei), respectively. An additional label ('unknown') was added for cells that had four or less similar neighbors. After accounting for this 'unknown' cell type, the overall accuracy was 99.3%. When applied to the study population, cell phenotype distributions showed significant differences across all diagnostic categories (Figure 3B). The proportion of Type 1 (mature) cells decreased with more advanced disease. In contrast, the proportions of Type 2 (small) and Type 3 (leukocytes) cells increased with disease progression. Median values for Type 1 (mature) and Type 2 (small) cells were significantly different between all lesion determinations. For Type 3 (leukocytes), all lesion determinations had significantly different median values except for benign vs. dysplasia ($p = 0.0539$).

The same cell identification model development process was completed on recently developed integrated instrumentation, cartridges, and cloud-based analysis tools. Images from two samples, one each from benign and malignant lesions, were collected with the POCOCT platform, and cell phenotype labels were overlaid on each recognized cell object (Figure 3C). Here, the benign lesion sample contained mostly Type 1 (mature) cells, while the malignant sample contained a mixture of primarily Type 2 (small), Type 3 (leukocytes), and Type 4 (lone nuclei).

Numerical Index and Diagnostic Models for Assessing PMOL

Expanding on this capability, a numerical index for discriminating benign and dysplasia/malignant lesions was developed using the cell phenotypes as predictors. Figure 4A shows the ROC curve representing discrimination performance of the multivariate model. The numerical index is a score between 0 and 100 that can be interpreted literally as the probability of dysplasia/malignancy. The diagnostic accuracy of the model is defined by the cutoff score that maximizes its AUC (benign vs. dysplasia/malignant numerical index cutoff of 36). Predictors for the model were retained as follows: cell phenotype distributions (Types 1, 2, and 3), age, sex, smoking pack years (i.e., packs per day times years of smoking), lesion size (maximum diameter), clinical impression of lesion as lichen planus, and lesion color (white, red, or both)

(Figure 4B). Minimal differences were observed between training and test error (28% and 27% misclassification rate on the training and test sets, respectively) which suggests no evidence of overfitting. The numerical index showed significant differences between all lesion diagnostic categories studied ($p < 0.01$) except for mild vs. moderate/severe dysplasia ($p = 0.1519$) (Figure 4C); however, significant differences were observed in a dichotomous model for mild vs. moderate dysplasia (i.e., 3|4) ($p = 0.04$). Model calibration shows the numerical index relative to the observed proportions of dysplasia/malignant subjects when sorted and grouped into deciles (Figure 4D). A non-significant result of the Hosmer-Lemeshow goodness of fit test suggests that there is no evidence of a poor fit ($p = 0.6259$).

Models were also developed for dichotomous classification across the OED spectrum, and Figure 5 summarizes the diagnostic performance of these models. The clinical algorithms resulted in AUCs ranging 0.81 (95% CI 0.76–0.86) for benign vs. mild dysplasia (3|4) to 0.97 (0.94–1.00) for healthy control (no lesion) vs. malignancy (0|6). While previous work demonstrated AUCs of 0.836 for the binary low vs. high risk (4|4) split and 0.883 for moderate vs. severe dysplasia (4|5),[11] these new optimized models here presented resulted in improved AUCs of 0.88 (0.84–0.93) and 0.92 (0.88–0.96) for the same diagnostic splits, respectively.

Cytopathology Software

A cytopathology interface tool was developed to assist pathologists in reviewing the brush cytology test results, enabling rich content cellular analyses on single- and multi-cell levels (Figure 6 and **Supplemental Figures**). This interface enables the pathologist users to access data stored and processed on cloud-based services, view results summaries, explore cytology results through data visualization tools, and generate automated oral cytopathology reports (Figure 7) which provide the adjunctive referral recommendations and summarize important information from cytology, including total cell count, cell phenotype distributions (Types 1, 2, and 3), and mean values for NC ratio, molecular biomarker fluorescence intensity, and cell circularity. The ability to assess cumulative data on this cloud-based cytopathology platform may improve pathologist decision making (e.g., through learning about their own histopathologic assessment vs. the POCOCT and, ultimately, the surgical pathology).

Discussion

This work demonstrates an evolution of the POCOCT technology towards a rapid and simple brush cytology analysis for POC or in a remote laboratory setting. We have demonstrated that (1) cell phenotypes can be accurately determined through the automated cytological assay and machine learning approach; (2) significant differences in cell phenotype distributions across diagnostic categories are found in three phenotypes (Types 1, 2, and 3); and (3) these cell phenotypes are valuable predictors for distinguishing lesion diagnostic categories in a multivariate lasso logistic regression model. The compilation of these results suggests that the observed cellular phenotypic variations within cytological samples are equated with disease severity and, thus, may be useful in the evaluation of PMOLs. Although cell phenotyping can be completed by a pathologist by manually identifying cells in a cytological sample, this is a lengthy process subject to human errors. Providing a means to automate metrics, such as the distributions of cell phenotypes, may increase adoption of this POCOCT approach through a cytopathology service and allow for pathologists to complete more efficient and more effective recommendations.

The optimized numerical index for evaluating PMOLs developed here represents a simple, practical, and effective approach that is directly applicable to clinical implementation and interpretation. While previous models relied on complicated high-dimensional cytological parameters, the classification and quantitation of cell phenotypes greatly simplifies the predictive algorithm and its interpretation, substantially improves performance for diagnostic splits relative to these earlier efforts,[11,14] and supports the translation of research methodologies from laboratory-based microscopy stations to an integrated POC instrument. With a total of 9 predictors, the practical model developed here represents a sparse solution (i.e., reduction of over 150 variables to 9) with greater potential generalizability without sacrificing any diagnostic performance. Further, excellent model calibration performance and significant differences between the diagnostic endpoints demonstrates strong potential for the numerical index as a continuous indicator of PMOL risk. While previous work was primarily focused on delivering binary results for referral decisions,[11] this new work involves a cytopathology interface tool, developed to assist pathologists in reviewing the brush cytology test results, and a numerical index, enabling rich content cellular analyses on single- and multi-cell levels. This interface enables the pathologist to access data stored on cloud-based services, view results summaries, explore cytology data through data visualization tools, and generate a

report that provides recommendations. Accurate diagnostic models spanning the entire OED spectrum also demonstrate the potential for the POCOCT to be used for multiple applications, such as screening PMOLs in primary care and the surveillance of patients with a history of OED and OSCC in secondary or tertiary care settings.

Although light-based adjuncts offer clinicians a new perspective to view a lesion at the POC, their diagnostic utility remains unproven.[5] Rashid and Warnakulasuriya reviewed the performance of light-based adjuncts in discriminating low and high risk lesions (VELscope [sensitivity/specificity: 30–100 / 15–100], ViziLite Plus [0–100 / 0–78], and Microlux DL [78 / 71]) and concluded that there is insufficient evidence to validate their efficacy as screening adjuncts.[31] Despite the numerous adjunctive tests available to assist in the diagnosis of PMOLs today, only cytology shows potential as a surrogate for gold standard histopathology.[32] Several commercial cytopathology services exist today including OralCDx (CDx Diagnostics, Inc.), OralCyte (ClearCyte Diagnostics, Inc.), Cyt ID (Forward Science), and ClearPrep OC (Resolution Biomedical). OralCDx, for example, provides an oral brush sample collection kit for their BrushTest.[33] Despite the ease of collection, samples need to be shipped to a commercial laboratory for analysis, resulting in delays between sample collection and test results. Further, the test often returns an ambiguous “atypical” result for which the positive predictive value for dysplasia or carcinoma has been determined to be only 30-40%.[34] Additionally, prior studies of cytology adjuncts demonstrated methodological gaps by only performing matched gold-standard histopathology on a subset of lesions with a higher index of suspicion for malignancy, and not for lesions with a lower index of suspicion which are frequently encountered in primary care settings.[35,36] A clinically validated POC cytology service capable of distinguishing the degree of OED in PMOL and stratifying the risk of malignant progression as a numerical index in near real-time would fulfill a significant unmet need mitigating unnecessary referrals to experts, leading to a more efficient process in surveillance clinics and reducing the patient distress related to waiting for test results.

One limitation is that previous studies of the POCOCT, and cytology adjuncts in general, primarily focused on PMOL evaluation in secondary care settings where the prevalence of dysplastic and malignant lesions may be substantially higher than in the primary care. Additionally, while expert clinicians in secondary and tertiary care settings have extensive training and experience in the recognition and risk

stratification of PMOLs, primary care clinicians may have difficulty distinguishing PMOLs from normal/non-neoplastic lesions. Thus, the POCOCT technology may potentially have a larger impact in primary care settings where there is a strong need to accurately interrogate the PMOLs detected there and generate a dichotomous outcome to indicate if referral of patients to higher care settings for expert evaluation and possible biopsy is required and if such referral should be urgent.

This manuscript provides a key step towards the development of new tools that could pave the way for new capabilities in the area of 'precision lesion diagnostics'. Helping to push forward this theme, we have demonstrated the utility of temporal changes in numerical index in a pilot study of Fanconi Anemia (FA) patients.[13] These efforts showed strong potential for patient-specific temporal changes in the lesion numerical index to track early signs of disease for this high risk population. Plans are now in place to (1) evaluate the POCOCT's precision lesion diagnostic capabilities through a prospective longitudinal study of malignant transformation and cancer recurrence and (2) move the POCOCT into a clinical trial to assess the POCOCT's diagnostic performance vs. routine care in primary care clinics.

Conclusion

In summary, we have demonstrated the utility of a POC-amenable cytology platform that has the potential to screen and monitor oral lesions across the entire diagnostic spectrum of OED. Cell phenotype distributions provided additional information in the assessment of PMOL. Further, a practical model comprised of patient information, lesion characteristics, and cell types from cytology showed similar performance characteristics to more complicated models previously developed. Cytopathology software may assist expert pathologists and non-expert care providers in reviewing and understanding the brush cytology test results. We developed data visualization tools to provide high content cellular analyses on single- and multi-cell levels with full transparency of test results data for pathologists. Additionally, oral cytopathology results summarize the test's most important predictors through indications of potential lesion progression for care providers and patients. Along with recently developed instrumentation and cartridges, this simple and sensitive system could provide non-invasive triage for PMOLs detected in primary, secondary, and tertiary care settings.

Future work may expand the utilization of molecular biomarkers and explore the identification of additional rare cell phenotypes to further improve performance. Future clinical studies may also be directed to determine: whether brush cytology could enable a wider sampling of large/multifocal lesion areas relative to incisional biopsies via multiple site-precise samplings; the effect of inflammation on the cytological analysis; whether the system can identify candida and distinguish clinical leukoplakia from neoplastic vs. non-neoplastic conditions; its placement in existing monitoring algorithms for PMOLs. Clinical trials are needed to assess the POCOCT's ability to identify early stage cancer relative to existing protocols and to validate the POCOCT as a substitute for biopsy. Future publications will describe and validate the integrated POC hardware (i.e., instrument, cartridge, and assay). To accelerate the translation and expand the adoption of the POCOCT platform, a cytopathology service for secondary and tertiary care oral cytology applications is now in development. Scaling and distribution of this versatile cytology approach is now underway with potential to serve diagnostic and surveillance applications in primary, secondary, and tertiary care settings.

References

1. Shield KD, Ferlay J, Jemal A, et al. The global incidence of lip, oral cavity, and pharyngeal cancers by subsite in 2012. *CA Cancer J Clin.* 2017;67(1):51-64.
2. National Cancer Institute Surveillance, Epidemiology, and End Results Program. Cancer stat facts: oral cancer and pharynx cancer. https://seer.cancer.gov/statfacts/html/oralcav.html. Accessed May 10, 2019.
3. Neville BW, Damm DD, Allen CM, Chi AC. Epithelial Pathology. In: Neville BW, Damm DD, Allen CM, Chi AC, eds. *J Oral Maxillofac Pathol.* 4th ed. St. Louis, MO: Elsevier Health Sciences; 2015:331-421.
4. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin.* 2018;68(1):7-30.
5. Huber MA. Adjunctive Diagnostic Techniques for Oral and Oropharyngeal Cancer Discovery. *Dent Clin North Am.* 2018;62(1):59-75.
6. Lingen MW, Abt E, Agrawal N, et al. Evidence-based clinical practice guideline for the evaluation of potentially malignant disorders in the oral cavity: A report of the American Dental Association. *J Am Dent Assoc.* 2017;148(10):712-727.
7. Lee J-J, Hung H-C, Cheng S-J, et al. Factors associated with underdiagnosis from incisional biopsy of oral leukoplakic lesions. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod.* 2007;104(2):217-225.
8. Weigum SE, Floriano PN, Christodoulides N, McDevitt JT. Cell-based sensor for analysis of EGFR biomarker expression in oral cancer. *Lab Chip.* 2007;7(8):995-1003.
9. Weigum SE, Floriano PN, Redding SW, et al. Nano-bio-chip sensor platform for examination of oral exfoliative cytology. *Cancer Prev Res.* 2010;3(4):518-528.
10. McDevitt JT, Weigum SE, Floriano PN, et al. A new bio-nanochip sensor aids oral cancer detection. *SPIE Newsroom.* 2011.
11. Abram TJ, Floriano PN, Christodoulides N, et al. 'Cytology-on-a-chip' based sensors for monitoring of potentially malignant oral lesions. *Oral Oncol.* 2016;60:103-111.

12. Speight PM, Abram TJ, Floriano PN, et al. Interobserver agreement in dysplasia grading: toward an enhanced gold standard for clinical pathology trials. *Oral Surg Oral Med Oral Pathol Oral Radiol.* 2015;120(4):474-482.

13. Abram TJ, Pickering CR, Lang AK, et al. Risk stratification of oral potentially malignant disorders in Fanconi Anemia patients using autofluorescence imaging and cytology-on-a chip assay. *Transl Oncol.* 2018;11(2):477-486.

14. Abram TJ, Floriano PN, James R, et al. Development of a cytology-based multivariate analytical risk index for oral cancer. *Oral Oncol.* 2019;92:6-11.

15. Bosman FT. Dysplasia classification: pathology in disgrace? *J Pathol.* 2001;194(2):143-144.

16. Warnakulasuriya S, Reibel J, Bouquot J, Dabelsteen E. Oral epithelial dysplasia classification systems: predictive value, utility, weaknesses and scope for improvement. *J Oral Pathol Med.* 2008;37(3):127-133.

17. Vigneswaran N, Beckers S, Waigel S, et al. Increased EMMPRIN (CD 147) expression during oral carcinogenesis. *Exp Mol Pathol.* 2006;80(2):147-159.

18. Torres-Rendon A, Roy S, Craig GT, Speight PM. Expression of Mcm2, geminin and Ki67 in normal oral mucosa, oral epithelial dysplasias and their corresponding squamous-cell carcinomas. *Br J Cancer.* 2009;100(7):1128-1134.

19. Jolliffe I. *Principal component analysis.* 2nd ed. New York: Springer; 2011.

20. Lugli E, Pinti M, Nasi M, et al. Subject classification obtained by cluster analysis and principal component analysis applied to flow cytometric data. *Cytometry A.* 2007;71A(5):334-344.

21. Hosmer DW, Lemeshow S. *Applied Logistic Regression.* 2nd ed. New York: John Wiley & Sons, Inc.; 2004.

22. LaValley MP. Logistic Regression. *Circulation.* 2008;117(18):2395-2399.

23. Hastie T, Tibshirani R, Friedman JH. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* 2nd ed. New York: Springer; 2009.

24. Wang D, Zhang W, Bakhai A. Comparison of Bayesian model averaging and stepwise methods for model selection in logistic regression. *Stat Med.* 2004;23(22):3451-3467.

25. Carpenter AE, Jones TR, Lamprecht MR, et al. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* 2006;7(10):R100.

26. El-Naggar AK, Chan JK, Grandis JR, Takata T, Slootweg PJ. *WHO classification of tumours of the head and neck.* 4th ed. Lyon: IARC Press; 2017.

27. Bouquot JE. Common oral lesions found during a mass screening examination. *J Am Dent Assoc.* 1986;112(1):50-57.

28. Sperandio M, Brown AL, Lock C, et al. Predictive value of dysplasia grading and DNA ploidy in malignant transformation of oral potentially malignant disorders. *Cancer Prev Res.* 2013;6(8):822-831.

29. Brands MT, Smeekens EAJ, Takes RP, et al. Time patterns of recurrence and second primary tumors in a large cohort of patients treated for oral cavity cancer. *Cancer Med.* 2019;8(12):5810-5819.

30. Mehanna HM, Rattay T, Smith J, McConkey CC. Treatment and follow-up of oral dysplasia — A systematic review and meta-analysis. *Head Neck.* 2009;31(12):1600-1609.

31. Rashid A, Warnakulasuriya S. The use of light-based (optical) detection systems as adjuncts in the detection of oral cancer and oral potentially malignant disorders: a systematic review. *J Oral Pathol Med.* 2015;44(5):307-328.

32. Lingen MW, Tampi MP, Urquhart O, et al. Adjuncts for the evaluation of potentially malignant disorders in the oral cavity: Diagnostic test accuracy systematic review and meta-analysis-a report of the American Dental Association. *J Am Dent Assoc.* 2017;148(11):797-813.e752.

33. CDx Diagnostics: The Painless Test for Common Oral Spots https://www.cdxdiagnostics.com/brushtest/. Accessed May 10, 2019.

34. Svirsky JA, Burns JC, Carpenter WM, et al. Comparison of computer-assisted brush biopsy results with follow up scalpel biopsy and histology. *Gen Dent.* 2002;50(6):500-503.

35. Sciubba JJ. Improving detection of precancerous and cancerous oral lesions: computer-assisted analysis of the oral brush biopsy. *J Am Dent Assoc.* 1999;130(10):1445-1457.

36. Poate TWJ, Buchanan JAG, Hodgson TA, et al. An audit of the efficacy of the oral brush biopsy technique in a specialist Oral Medicine unit. *Oral Oncol.* 2004;40(8):829-834.

Figure Legends

Figure 1. Diagnostic categories for oral cancer and dysplasia based on WHO classification with 5-year malignant transformations and 5-year cancer recurrence rates. While 10% of US adults may present to their dentist for a routine care visit with an abnormal oral cavity lesion, about 83% of these lesions are diagnosed clinically as having no malignant potential, and 17% have unknown significance and meet the clinical criteria for PMOL. About 17% of PMOLs are histopathologically diagnosed with OED or OSCC. OED is about 15 times more common than OSCC, yet only a fraction of patients with dysplastic PMOLs undergo malignant transformation.

Figure 2. The POCOCT assay platform allows for the analysis of cellular samples obtained from a minimally invasive brush cytology sample. The cell suspension collected in this manner allow for the simultaneous quantification of cell morphometric data and expression of molecular biomarkers of malignant potential in an automated manner using refined image analysis algorithms based on pattern recognition techniques and advanced statistical methods. This novel approach turns around cytology results in a matter of minutes as compared to days for traditional pathology methods, thereby making it amenable to POC settings. The POC testing is expected to have tremendous implications for disease management by enabling dental practitioners and primary care physicians to circumvent the need for multiple referrals and consultations before obtaining assessment of molecular risk of PMOL.

Figure 3. A cell type identification model was developed to automatically classify cell Types 1-4. Panel A (left) shows the four distinct cell phenotypes that were identified: Type 1 (‘mature squamous cells’), Type 2 (‘small round cells’), Type 3 (‘leukocytes’), and Type 4 (‘lone nuclei’). Principal component analysis (right) shows cell phenotypes clustered into distinct groups with substantial separation between cell phenotype labels, demonstrating strong promise for an effective cell phenotype recognition algorithm. Boxplots in Panel B show the study population distributions of mature squamous cells (left), small round cells (center), and leukocytes (right), representing the predicted mean cell type percentages across six biomarker assays (αvβ6, CD-147, EGFR, geminin, Ki-67, and MCM2) within each lesion class: normal (n=121), benign (n=241), dysplasia (n=59), and malignant (n=65). The results shown include only patients with definitive

lesion determinations and patients with evaluable data for all six biomarkers. Panel C shows limited field of view cytology pseudocolor images (fluorescence images acquired with a monochrome camera and digitally assigned to red, green, and blue color channels) of benign (left) and malignant (right) lesions with the cell phenotype model output labels overlaid as follows: “M” for mature squamous cells, “S” for small round cells, “W” for leukocytes, and “L” for lone nuclei (Unknown type “U” not shown). Fluorescent staining shows the cytoplasm (red), nuclei (blue), and Ki-67 biomarker (green).

Figure 4. Algorithm results of the dichotomous benign vs. dysplasia/malignant lesion model from 241 benign lesion and 124 dysplasia and malignant lesion subjects for six molecular biomarker assays on the POCOCT system. Panel A shows the ROC curve for the model. The lasso logistic regression coefficients are provided in Panel B. The predictors are as follows: “1-%TYPE 1” (percent of cells that are non-mature squamous cells), “%TYPE 2” (percent of cells that are small round cells), “%TYPE 3” (percent of cells that are leukocytes), “AGE”, “SEX”, “PACKYR” (pack years), “LSIZEMAX” (lesion diameter of the major axis), “LICHENFN” (clinical impression of lichen planus), and “LESIONCOLOR” (red, white, or red/white). The boxplot in Panel C shows cross-validated algorithm response (“numerical index”) for the lasso logistic regression on the test set averaged over all biomarker assays. Distribution of scores are represented for benign (n=241), mild dysplasia (n=38), moderate/severe dysplasia (n=21), and malignant lesions (n=65). Panel D shows a model calibration plot of the predicted responses (numerical index) sorted and grouped into deciles vs. the observed proportions of dysplasia and malignant lesions.

Figure 5. Diagnostic models for the OED spectrum. Results are shown for the cross-validated clinical algorithms for benign vs. dysplasia (2|3), mild vs. moderate dysplasia (3|4), low vs. high risk (4|4), moderate vs. severe dysplasia (4|5), healthy control (no lesion) vs. malignant (0|6), and benign dysplasia vs. malignant (2|6) models. Model responses for each subject were averaged over all biomarker assays to inform diagnostic performance. AUC, sensitivity, and specificity are mean and 95% confidence interval values for the cross-validated test set.

Figure 6. Cytopathology interface tool provides pathologists with cloud access to test results summaries and detailed data visualizations (A), scatter plots (B), and histograms (C) for over 150 different cytology parameters. With this tool, pathologists can view all cells within the field of view, zoom in for more detail, and isolate individual cells of interest.

Figure 7. Oral cytopathology test results. The algorithm result is a numerical index between 0 and 100 with a cutoff of 36 that distinguishes benign and dysplasia/malignant ("atypical") lesions (left). Other informative cytopathology results are displayed on a reference range, including total cell counts, cell phenotype distributions, mean values for NC ratio, molecular biomarker fluorescence intensity, and cell circularity. Images and outlines of the cells are provided for additional test context (right).

Supplemental Figures

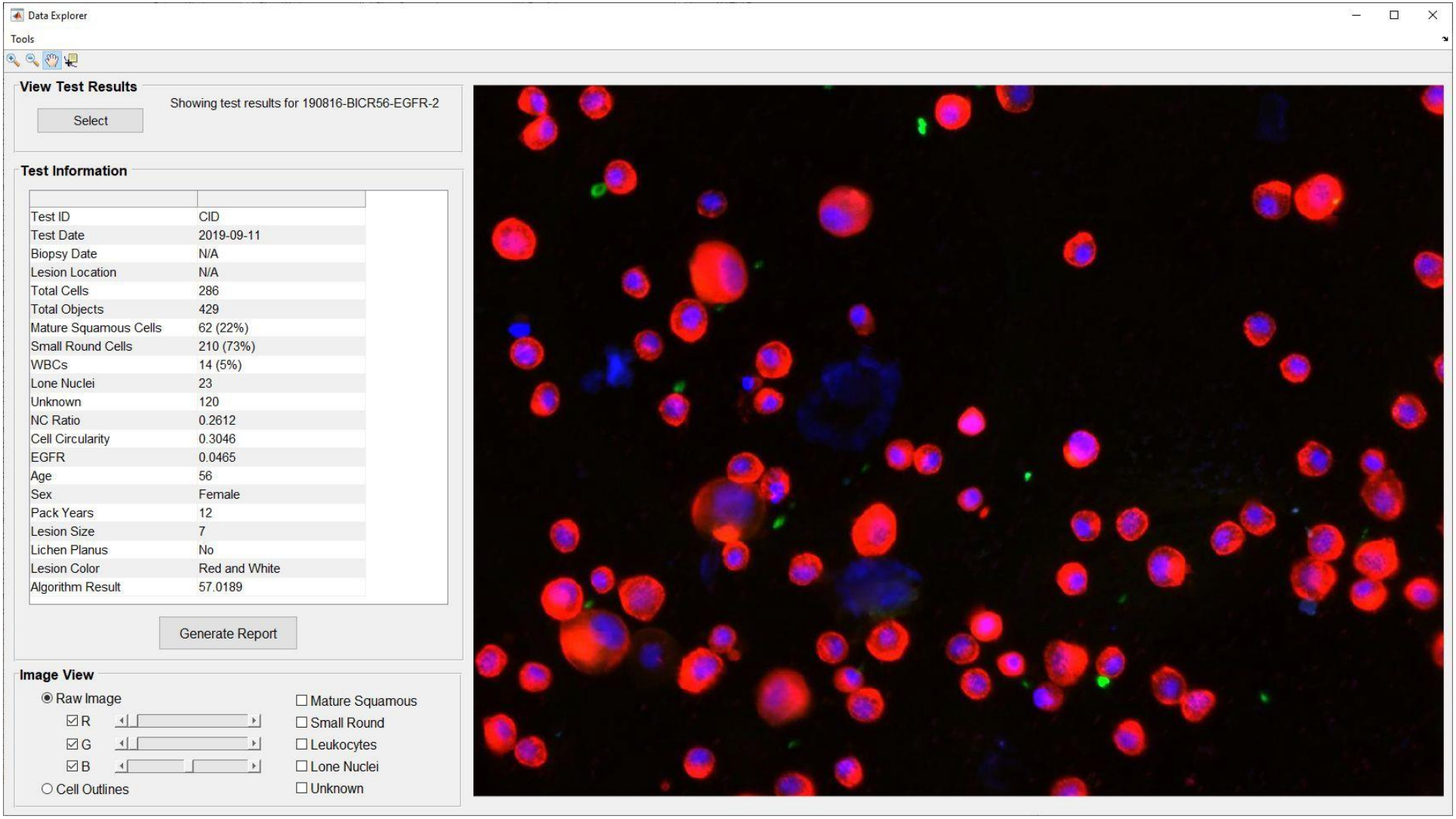

Figure S1. Screenshot of cytopathology interface showing BICR 56 cancer cells magnified view with all three fluorescent labels (red: phalloidin, green:

EGFR, blue: DAPI).

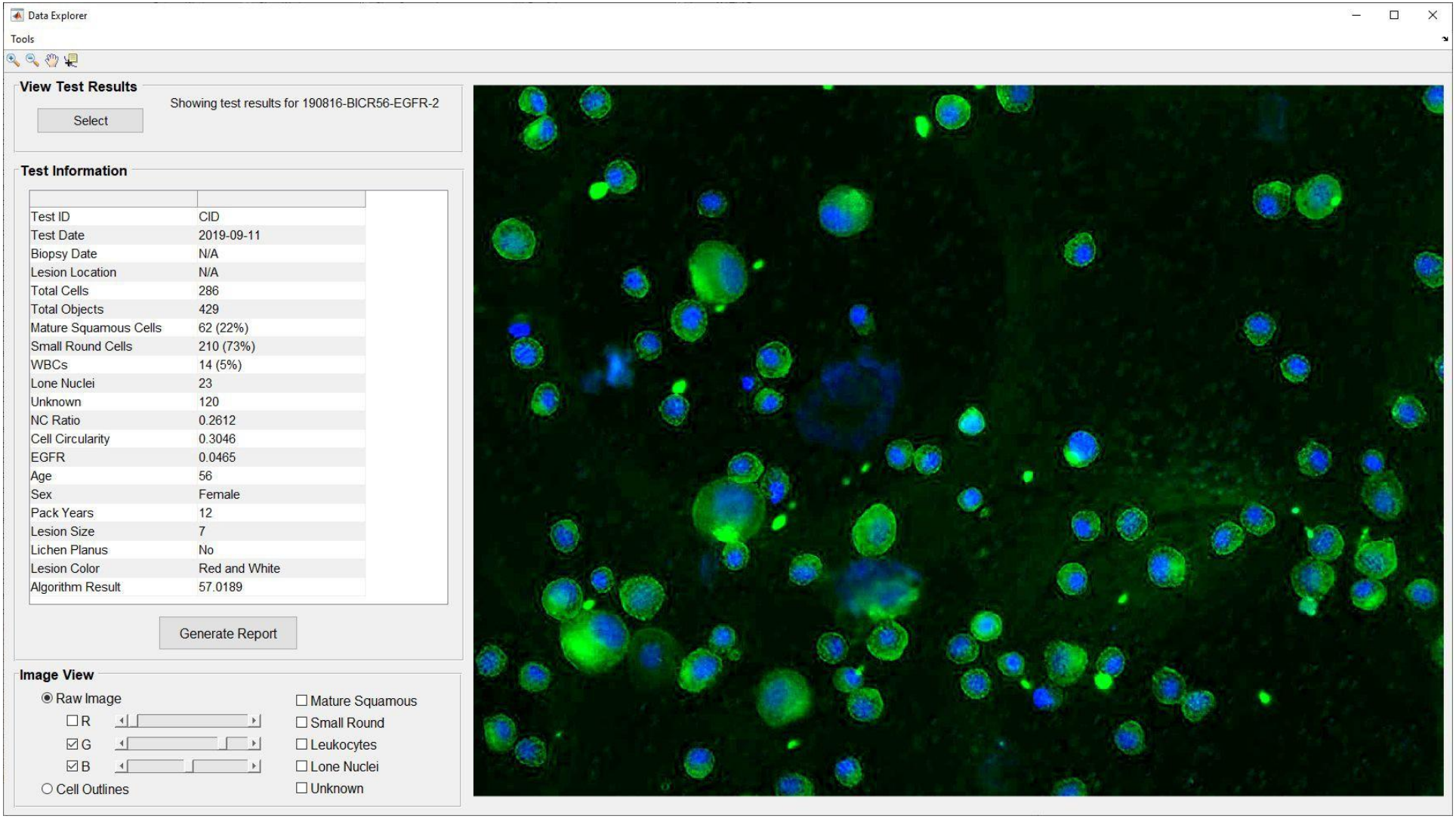

Figure S2. Screenshot of cytopathology interface showing BICR 56 cancer cells magnified view with green (EGFR) and blue (DAPI) fluorescent labels.

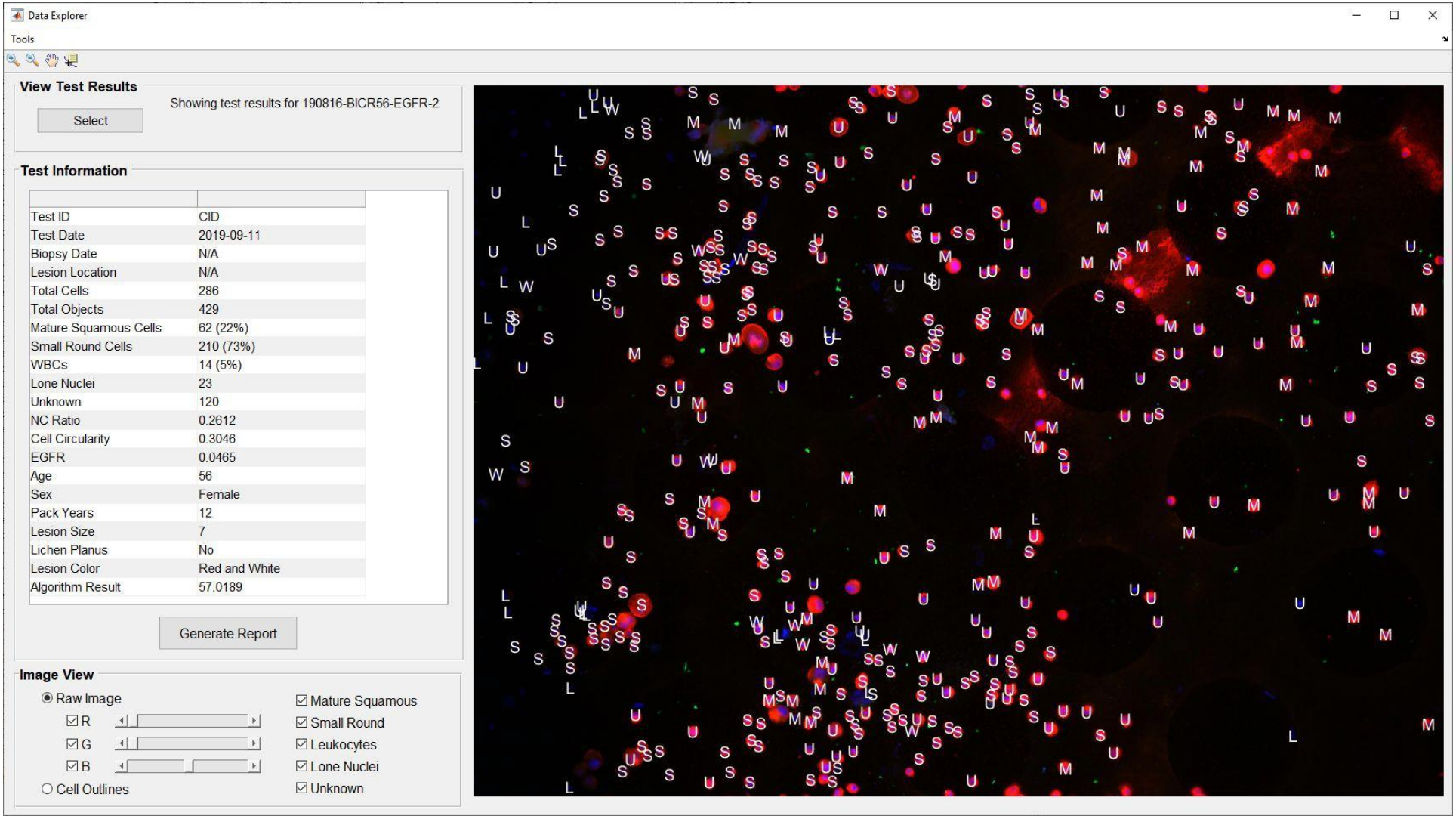

Figure S3. Screenshot of cytopathology interface showing BICR 56 cancer cells with cell phenotype labels overlaid (M: mature squamous, S: small

round, W: leukocytes, L: lone nuclei, U: unknown).

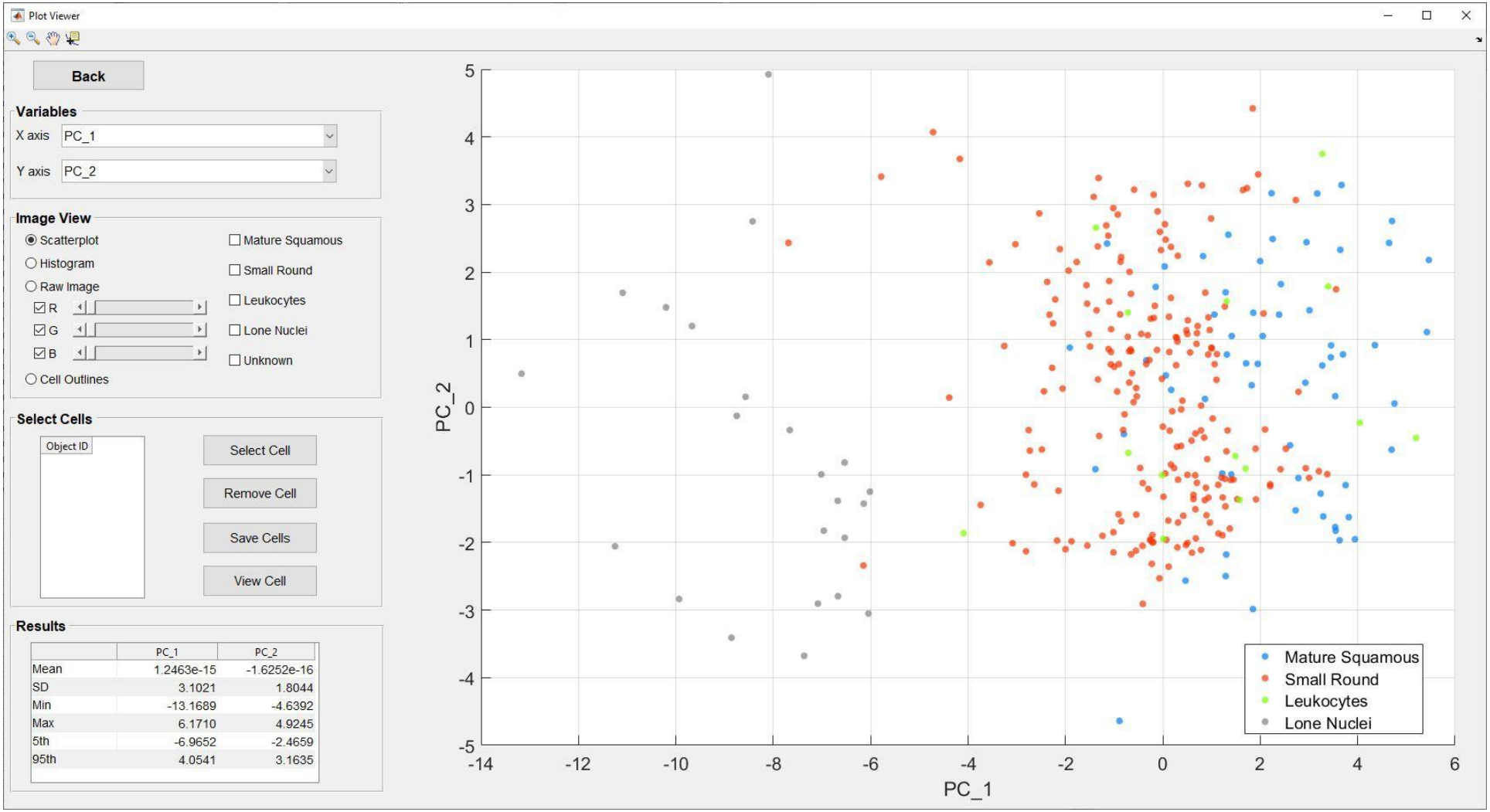

Figure S4. Screenshot of cytopathology interface showing a principal component scatter plot from a sample of BICR 56 cancer cells.

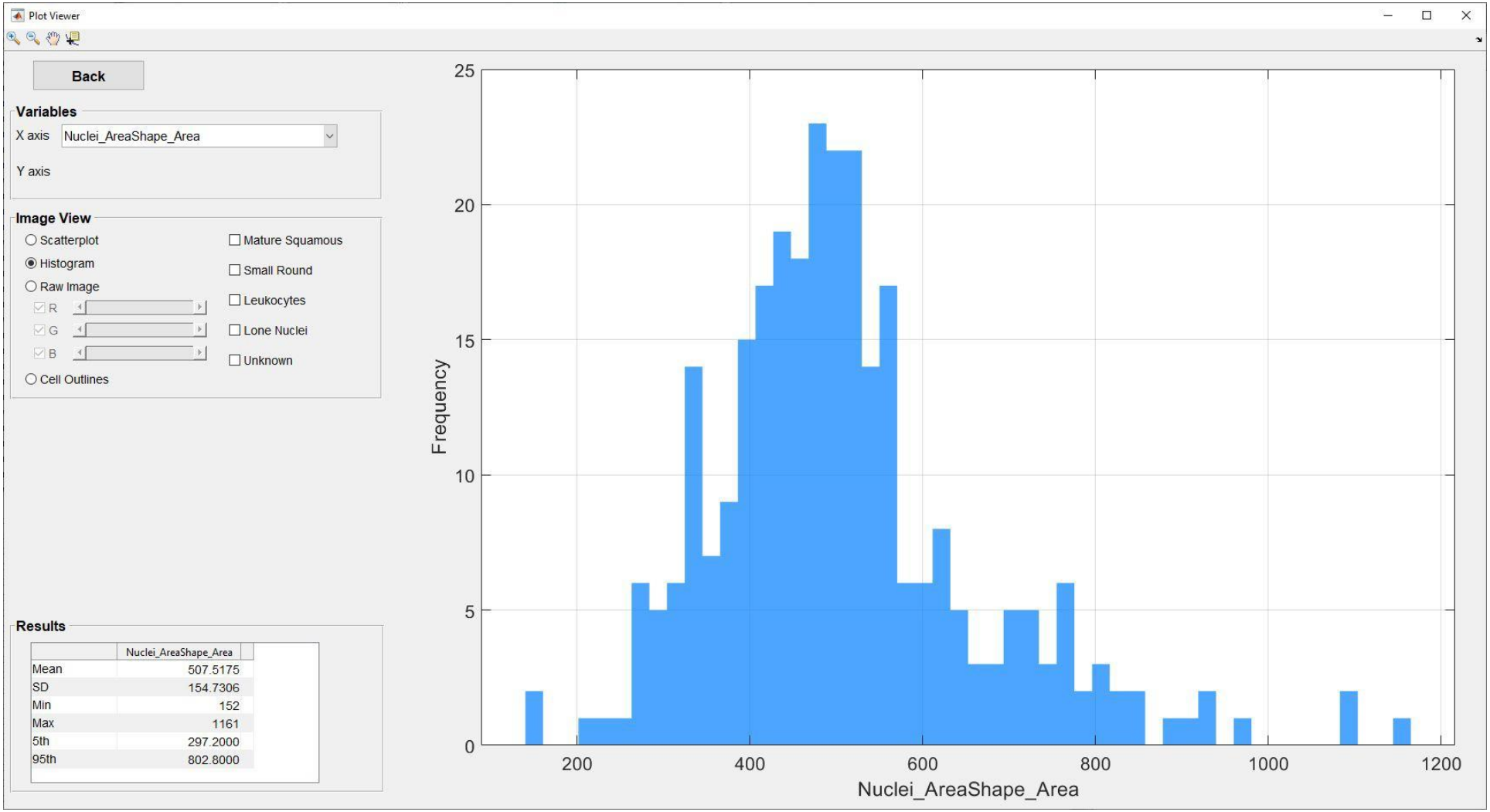

Figure S5. Screenshot of cytopathology interface showing histogram of nuclear area measurements from a sample of BICR 56 cancer cells.

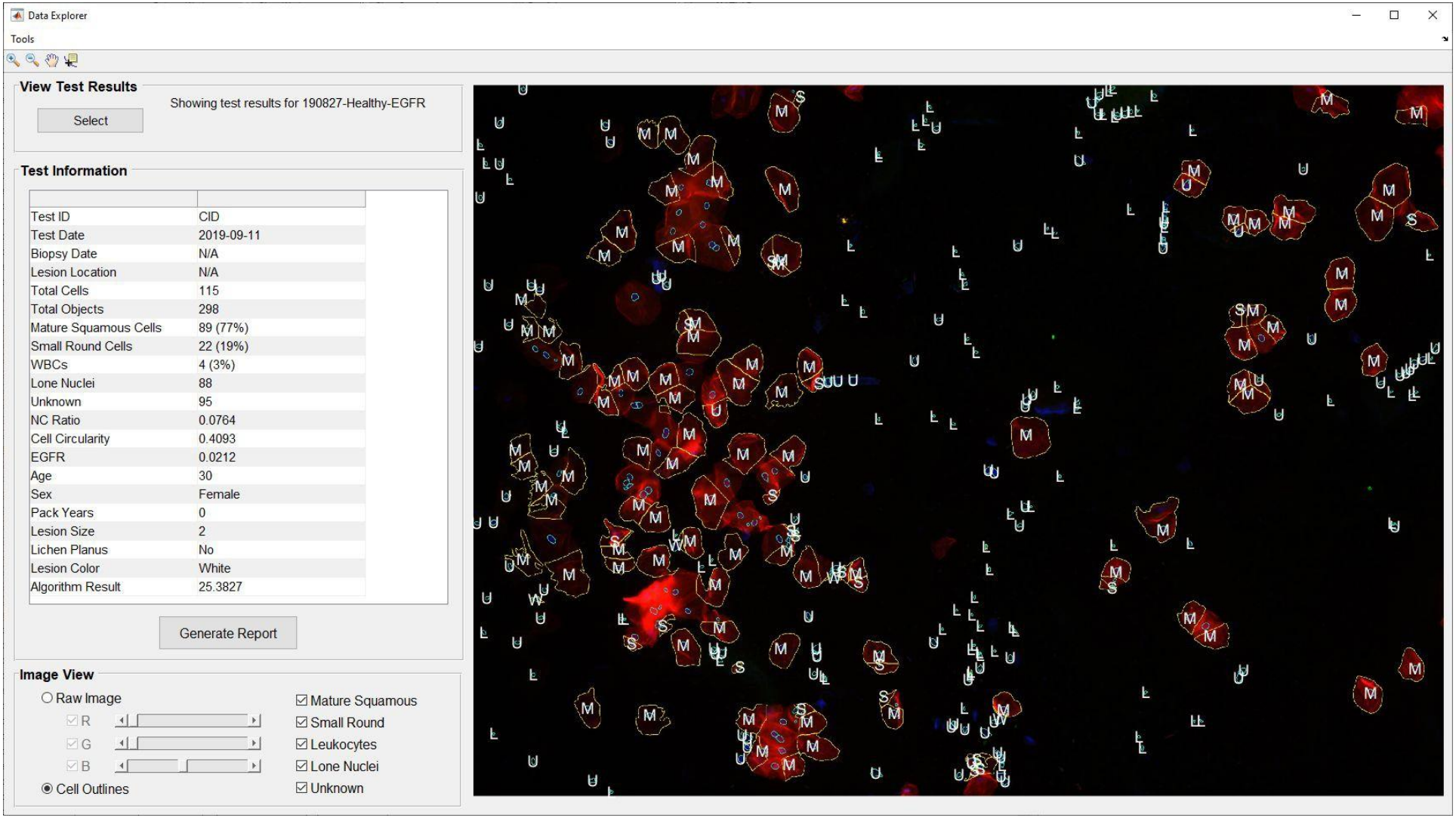

Figure S6. Screenshot of cytopathology interface showing brush biopsy sample of healthy control cells with cell phenotype labels overlaid (M: mature

37 squamous, S: small round, W: leukocytes, L: lone nuclei, U: unknown, not shown).

38
39
40
41
42
43
44
45
46
47

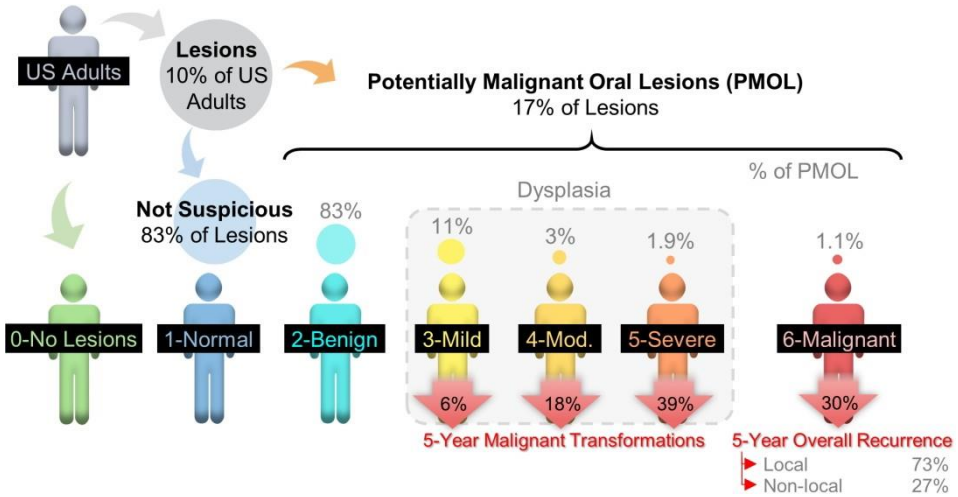

Figure 1. Diagnostic categories for oral cancer and dysplasia based on WHO classification with 5-year malignant transformations and 5-year cancer recurrence rates. While 10% of US adults may present to their dentist for a routine care visit with an abnormal oral cavity lesion, about 83% of these lesions are diagnosed clinically as having no malignant potential, and 17% have unknown significance and meet the clinical criteria for PMOL. About 17% of PMOLs are histopathologically diagnosed with OED or OSCC. OED is about 15 times more common than OSCC, yet only a fraction of patients with dysplastic PMOLs undergo malignant transformation.

171x101mm (300 x 300 DPI)

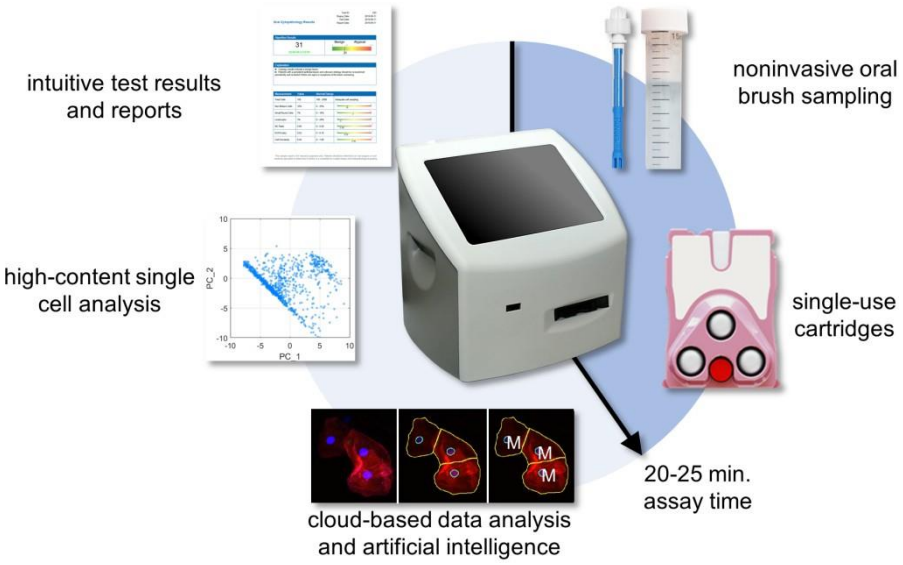

Figure 2. The POCOCT assay platform allows for the analysis of cellular samples obtained from a minimally invasive brush cytology sample. The cell suspension collected in this manner allow for the simultaneous quantification of cell morphometric data and expression of molecular biomarkers of malignant potential in an automated manner using refined image analysis algorithms based on pattern recognition techniques and advanced statistical methods. This novel approach turns around cytology results in a matter of minutes as compared to days for traditional pathology methods, thereby making it amenable to POC settings. The POC testing is expected to have tremendous implications for disease management by enabling dental practitioners and primary care physicians to circumvent the need for multiple referrals and consultations before obtaining assessment of molecular risk of PMOL.

171x96mm (300 x 300 DPI)

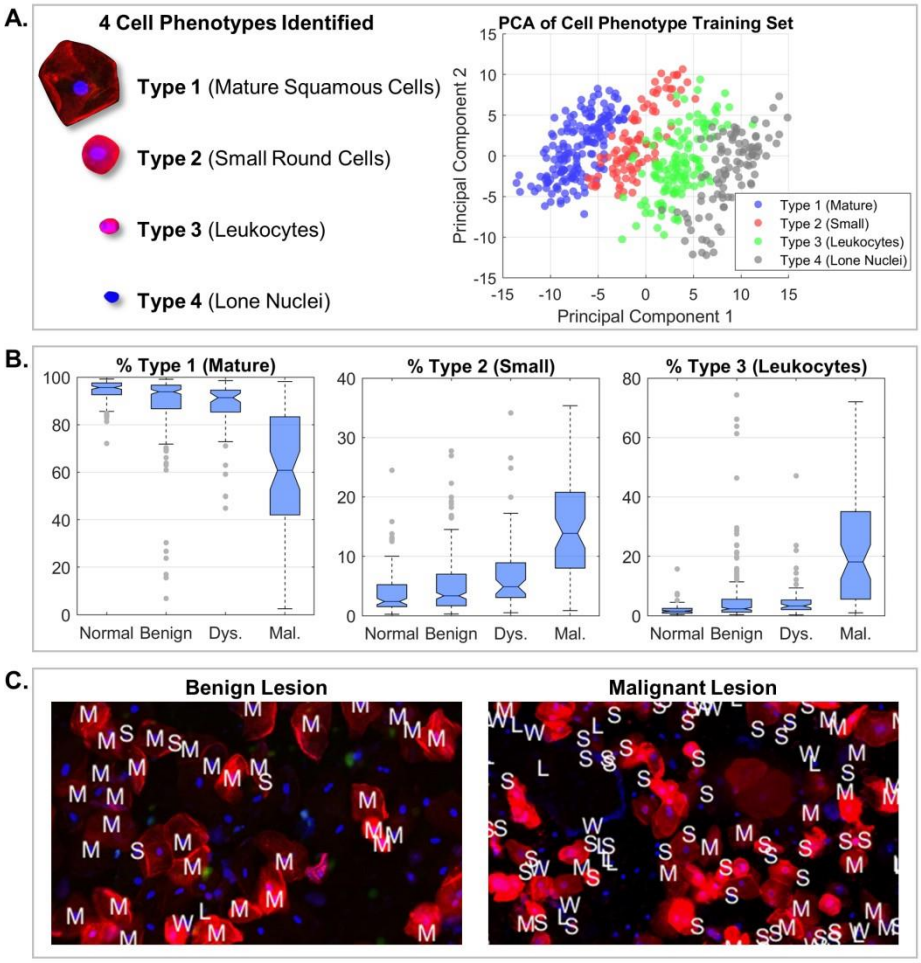

Figure 3. A cell type identification model was developed to automatically classify cell Types 1-4. Panel A (left) shows the four distinct cell phenotypes that were identified: Type 1 ('mature squamous cells'), Type 2 ('small round cells'), Type 3 ('leukocytes'), and Type 4 ('lone nuclei'). Principal component analysis (right) shows cell phenotypes clustered into distinct groups with substantial separation between cell phenotype labels, demonstrating strong promise for an effective cell phenotype recognition algorithm. Boxplots in Panel B show the study population distributions of mature squamous cells (left), small round cells (center), and leukocytes (right), representing the predicted mean cell type percentages across six biomarker assays (αvβ6, CD-147, EGFR, geminin, Ki-67, and MCM2) within each lesion class: normal (n=121), benign (n=241), dysplasia (n=59), and malignant (n=65). The results shown include only patients with definitive lesion determinations and patients with evaluable data for all six biomarkers. Panel C shows limited field of view cytology pseudocolor images (fluorescence images acquired with a monochrome camera and digitally assigned to red, green, and blue color channels) of benign (left) and malignant (right) lesions with the cell phenotype model output labels overlaid as follows: "M" for mature squamous cells, "S" for small round cells, "W" for leukocytes, and "L" for lone nuclei (Unknown type "U" not shown). Fluorescent staining shows the cytoplasm (red), nuclei (blue), and Ki-67 biomarker (green).

171x171mm (300 x 300 DPI)

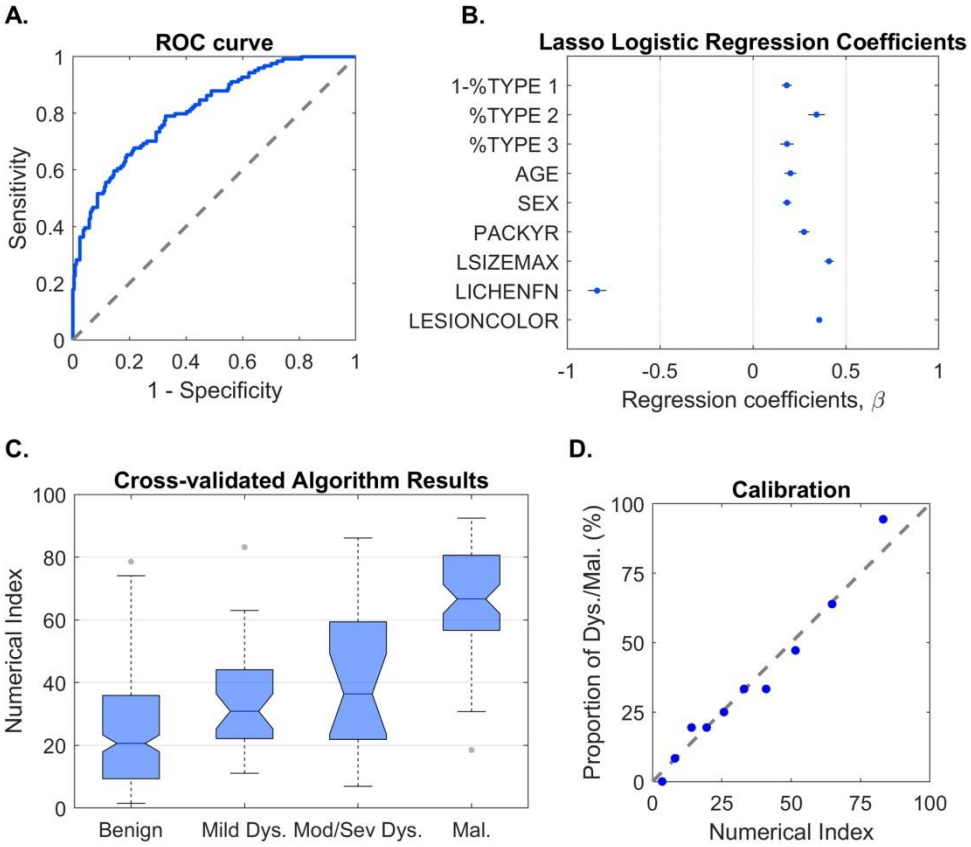

Figure 4. Algorithm results of the dichotomous benign vs. dysplasia/malignant lesion model from 241 benign lesion and 124 dysplasia and malignant lesion subjects for six molecular biomarker assays on the POCOCT system. Panel A shows the ROC curve for the model. The lasso logistic regression coefficients are provided in Panel B. The predictors are as follows: "1-%TYPE 1" (percent of cells that are non-mature squamous cells), "%TYPE 2" (percent of cells that are small round cells), "%TYPE 3" (percent of cells that are leukocytes), "AGE", "SEX", "PACKYR" (pack years), "LSIZEMAX" (lesion diameter of the major axis), "LICHENFN" (clinical impression of lichen planus), and "LESIONCOLOR" (red, white, or red/white). The boxplot in Panel C shows cross-validated algorithm response ("numerical index") for the lasso logistic regression on the test set averaged over all biomarker assays. Distribution of scores are represented for benign (n=241), mild dysplasia (n=38), moderate/severe dysplasia (n=21), and malignant lesions (n=65). Panel D shows a model calibration plot of the predicted responses (numerical index) sorted and grouped into deciles vs. the observed proportions of dysplasia and malignant lesions.

171x152mm (300 x 300 DPI)

Model	Non-case / Case	Sensitivity	Specificity	AUC
2|3	2 3 4 5 6	0.69 (0.64–0.74)	0.77 (0.72–0.81)	0.81 (0.76–0.86)
3|4	2 3 4 5 6	0.79 (0.74–0.83)	0.85 (0.81–0.89)	0.88 (0.84–0.93)
4|4	2 3 4 5 6	0.78 (0.73–0.82)	0.87 (0.83–0.90)	0.88 (0.84–0.93)
4|5	2 3 4 5 6	0.82 (0.78–0.86)	0.88 (0.84–0.91)	0.92 (0.88–0.96)
2|6	2 6	0.89 (0.85–0.92)	0.90 (0.85–0.93)	0.95 (0.91–0.98)
0|6	0 6	0.94 (0.89–0.97)	0.92 (0.87–0.95)	0.97 (0.94–1.00)

Figure 5. Diagnostic models for the OED spectrum. Results are shown for the cross-validated clinical algorithms for benign vs. dysplasia (2|3), mild vs. moderate dysplasia (3|4), low vs. high risk (4|4), moderate vs. severe dysplasia (4|5), healthy control (no lesion) vs. malignant (0|6), and benign dysplasia vs. malignant (2|6) models. Model responses for each subject were averaged over all biomarker assays to inform diagnostic performance. AUC, sensitivity, and specificity are mean and 95% confidence interval values for the cross-validated test set.

171x99mm (300 x 300 DPI)

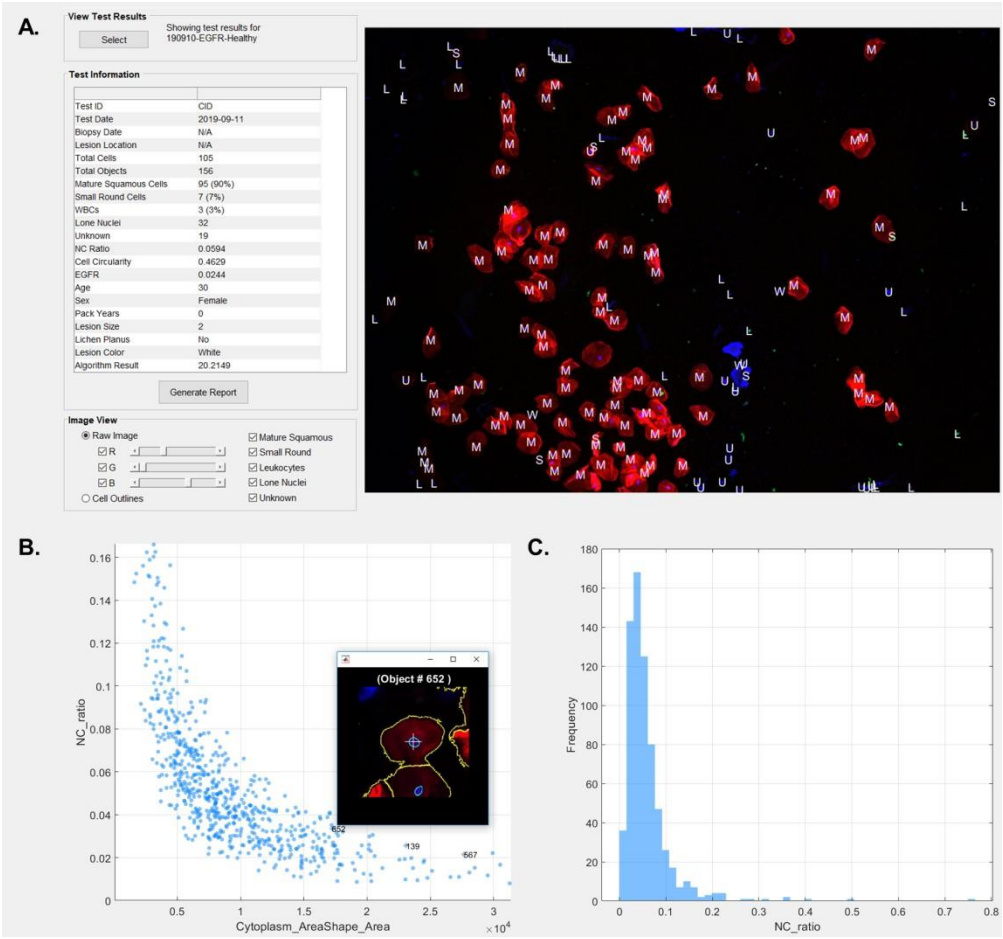

Figure 6. Cytopathology interface tool provides pathologists with cloud access to test results summaries and detailed data visualizations (A), scatter plots (B), and histograms (C) for over 150 different cytology parameters. With this tool, pathologists can view all cells within the field of view, zoom in for more detail, and isolate individual cells of interest.

171x160mm (300 x 300 DPI)

Figure 7. Oral cytopathology test results. The algorithm result is a numerical index between 0 and 100 with a cutoff of 36 that distinguishes benign and dysplasia/malignant ("atypical") lesions (left). Other informative cytopathology results are displayed on a reference range, including total cell counts, cell phenotype distributions, mean values for NC ratio, molecular biomarker fluorescence intensity, and cell circularity. Images and outlines of the cells are provided for additional test context (right).

171x106mm (300 x 300 DPI)

Supplemental Methods

Biomarker Selection Rationale

Six molecular biomarkers were selected (αvβ6, CD147, EGFR, geminin, Ki67, and MCM2) based on their capacity to distinguish benign, dysplastic, and malignant oral epithelial cells through prior immunohistochemistry studies.[1-3] These markers fall into three groups based on their localization: cell membrane, cytoplasm, and nucleus. **Table S1** summarizes the molecular biomarkers used in the study.

Table S1. Summary of molecular biomarkers

Biomarker	Localization	Function
αvβ6	CM	an integrin receptor undetectable in normal oral epithelium, but highly expressed in dysplasia and OSCC[4,5]
CD147	CM	a multifaceted molecule that facilitates tumor progression by several mechanisms[6]
EGFR	CM + C	a transmembrane glycoprotein whose overexpression may contribute to tumor progression[7]
Geminin	N + C	a marker of proliferation[2]
Ki67	N	a marker of proliferation that is overexpressed at initial stages of oral carcinogenesis[7]
MCM2	N	an essential component for DNA replication associated with deregulated expression in dysplastic and malignant epithelial cells[8,9]

* CM: cell membrane; C: cytoplasm; N: nucleus

Patient Recruitment

Data used in this study originated from the 999-patient multisite prospective non-interventional study evaluating the cytology-on-a-chip system for the measurement of cytological parameters on brush cytology samples to assist in the diagnosis of PMOL. Briefly, both histopathological and brush cytological samples for 714 subjects from three patient groups were measured: (1) subjects with PMOL who underwent scalpel biopsy as part of the standard of care for microscopic diagnosis, (2) subjects with recently diagnosed malignant lesions, and (3) healthy

volunteers without lesions. Only subjects with complete biomarker results were included in the analysis (N = 486). **Table S2** summarizes the patient characteristics of those subjects included in the analysis.

Table S2. Patient characteristics and histopathological diagnoses

Characteristics and Histopathological Diagnoses	N (%)
Total	486
Sex	
Male	211 (43.4)
Female	275 (56.6)
Age	
>60	165 (34.0)
≤60	320 (65.8)
Patient Group	
Healthy Volunteer	121 (24.9)
Subjects with Previously Diagnosed Malignant Lesion	36 (7.4)
Subject with a Potentially Malignant Lesion	329 (67.7)
Histopathological Diagnosis	
Normal	121 (24.9)
Benign	241 (49.6)
Mild Dysplasia	38 (7.8)
Moderate Dysplasia	12 (2.5)
Severe Dysplasia	9 (1.9)
Malignant	65 (13.4)

Clinical Protocol

The clinical protocol for this study was published previously[10] and is summarized as follows. Patients in group 1 underwent brush sampling of the oral lesion and a brush sampling of

the contralateral, clinically normal mucosa. The brush cytology sample was taken immediately

before the same lesion underwent a scalpel biopsy. Patients in group 2 underwent brush biopsy of the known cancerous lesion, as well as the contralateral, clinically normal mucosa. For healthy volunteers in group 3, a brush biopsy of normal appearing tissue on the lateral or ventral surface of the tongue and a brush biopsy of normal appearing tissue on the left or right buccal mucosa were taken. Brush biopsy samples were taken using a soft Rovers Orcellex oral cytology brush (Rovers Medical Devices B.V., Oss, The Netherlands). The brush was applied directly to the lesion or control oral mucosa using mild pressure and rotated 360 degrees approximately 10-15 times in the same direction to obtain the cytologic sample.

Cytology-on-a-Chip Protocol

The following methods have been published previously[11] and are summarized here for convenience. Immediately after brush cytology samples were collected, cells were harvested by vortexing the brush head in minimum essential medium (MEM) culture media, followed by a PBS wash, re-suspension in FBS containing 10% of the cryo-preservative dimethyl-sulfoxide (DMSO), frozen, and stored in a -80 degrees C freezer.

Prior to processing on the device, patient samples were thawed rapidly in a 37 degrees C water bath, washed with PBS, and fixed for one hour in 0.5% formaldehyde prepared fresh from a 16% stock solution (Polysciences, Warrington, PA, #18814-20). After fixation, cells were washed twice in PBS, re-suspended in 150 µL 0.1% PBS with 0.1% BSA (PBSA), and stored at 40 degrees C until ready to process. Before sample delivery, the cell suspension was diluted in a 20% glycerol/0.1% PBSA solution to improve cell distribution across the membrane and to reduce cell clumping.

Using a custom built manifold connecting external fluidic tubing to the inlet and outlet ports of the microfluidic device, the assembly was positioned on a robotically controlled microscope stage (ProScan II, Prior Scientific, Cambridge, UK) and connected to a peristaltic pump (SciQ 400, Watson Marlow, Wilmington, MA) and manually controlled 6-position injector valve (Vici,

Valco Instruments, Houston, TX). Antibody stock solutions were vortexed for 30 seconds and centrifuged at 14,000 rpm for 5 minutes before preparing working dilutions to avoid precipitates.

All assays contained Phalloidin and DAPI in the secondary antibody cocktail, but each was specific for a single molecular biomarker primary-secondary antibody pair. Working dilutions of antibodies were prepared in 0.1% PBSA with 0.1% Tween-20 (EMD Millipore, Billerica, MA, # 655206). Primary monoclonal antibodies were raised from either mouse (EGFR [Life Technologies, Carlsbad, CA, #MS-378-P, 10 µg/mL]), rabbit (αvβ6 [Abcam, Cambridge, MA, #Ab124968, 6 µg/mL], Ki67 [Abcam #Ab15580, 29 µg/mL], and MCM2 [Abcam #Ab108935, 10 µg/mL]), or goat (CD-147 [EMMPRIN] [R&D Systems, Minneapolis, MN, #AF972, 20 µg/mL]. AlexaFluor-488 conjugated secondary antibodies were specific for F $(ab')_2$ fragments of mouse IgG (Life Technologies #A11017, 20 µg/mL for EFGR), rabbit IgG (Life Technologies #A11070, 50 µg/mL for αvβ6, 64 µg/mL for Ki67, and 23.5 µg/mL for MCM2), or goat IgG (Life Technologies #A11078, 40 µg/mL for CD147). A working concentration of 0.33 µM was used for Phalloidin-AlexaFluor-647 (Life Technologies #A22287) and 5 µM for DAPI (Life Technologies #D3571).

In summary, the lab-on-a-chip sample processing was comprised of the following steps: 1) the device was primed with PBS at a flow rate of 735 µL/min for 2 minutes, 2) the cell suspension in 20% glycerol/0.1% PBSA was delivered at 1.5 mL/min for 2 minutes, 3) cells were washed with PBS at 1 mL/min for 2.5 min, 4) the primary antibody solution was delivered through a 0.2 µm PVDF syringe filter at 250 µL/min for 2.5 min, 5) a wash step similar to step 3 was performed, 6) the secondary antibody solution was delivered under the same conditions as step 4, 7) a final wash step was performed, and 8) automated image capture was performed.

Sample Digitization

More complete details on cytology sample digitization and a complete list of intensity and morphological parameters can be found in our previous publication.[11] Images were recorded with a motorized reflected fluorescence microscope (Olympus BX-RFAA) equipped with a CCD

camera (Hamamatsu ORCA-03G) through a 10x objective (10x/0.30NA UPlanFl, Olympus). A total of 25 unique fields of view (FOVs) repeated for 3 different z-focal planes were automatically captured across a 20 mm^2 area using a robotic x-y-z microscope stage. Due to the complex three-dimensional morphology of oral squamous cells, multiple z-focal planes were captured and subsequently combined into a single, enhanced depth-of-field image to simplify the multi-spectral detection of the three fluorescent labels using ImageJ "stack focuser".

Combinations of custom macros and the open-source image analysis tools ImageJ[12] and Cell Profiler[13] were developed to automatically detect individual cells and define their nuclear and cytoplasmic boundaries as individual regions of interest (ROI). These ROIs were used to obtain intensity measurements associated with the three spectral channels and were used to define morphometric parameters. The DAPI and Phalloidin molecular labels served primarily to assist in the automated segmentation of individual nuclei and cytoplasm, respectively.

Cell Identification Model Training and Validation

A cell type classification model was explored for its ability to discriminate and quantitate the frequency and distributions of four cell types: Type 1 (mature squamous cells), Type 2 (small round cells), Type 3 (leukocytes), and Type 4 (lone nuclei). To recognize these phenotypes, a machine learning algorithm was trained on 144 cellular/nuclear features from single-cell analyses, including morphological and intensity-based measurements. A training set was manually compiled by randomly selecting and labeling cells, resulting in approximately 100-200 single-cell objects for each of the four cell types. All features were log-normalized and standardized for zero mean and unit variance. Principal component analysis (PCA) was performed on the training set, and a scatterplot of the first two principal components was generated to visualize the internal data structure and variance. A *k*-nearest neighbors (*k*-NN) classifier was trained on the standardized features using 10-fold cross-validation and configured to find the nearest 7 neighbors in feature space (Euclidean distance). Cross-validated predicted responses by the *k*-NN classifier were

recorded, and accuracy was reported for the overall cross-validation set and individually for each of the four cell types. *k*-NN model responses with 4 or less out of 7 similar neighbors were labelled "unknown" type, and cross-validated accuracy was reported for the overall training set after accounting for unknown object types.

The cell type classification model was retrained on the entire training dataset, and this final model was applied to the study population and averaged across each of the six molecular biomarker assays. Results are presented for only subjects with evaluable data for all biomarker measurements (N = 486). Boxplots were generated to show the distributions of cell phenotypes across 4 diagnostic categories as follows: 121 normal/non-neoplastic, 241 benign, 59 dysplasia, and 65 malignant. Median values of cell phenotypes were compared for all lesion determinations using a two-sided Wilcoxon rank sum test at a significance level of $p = 0.05$. Cell phenotype frequencies and distributions for each subject were retained for use in clinical algorithm development.

The same cell type identification model development process was completed on recently developed integrated instrument, cartridges, and cloud-based analysis tools. Images of benign and malignant lesions were collected with this cloud POC cytology platform, and cell phenotype labels were overlaid on each recognized cell object.

Numerical Index and Diagnostic Models for Assessing PMOL

The analysis of dichotomous outcomes with mutually exclusive levels is common in clinical diagnostics, and logistic regression is regarded as the standard method of analysis for these situations attributed to its probabilistic interpretation and ability to function as a dichotomous classifier. Clinical data are often challenged by high-dimensionality and highly correlated predictors that may generate model coefficients with high variance. For these situations, a size penalty as implemented by the lasso technique may be applied to shrink the effect sizes and reduce coefficient variability. Additionally, the lasso technique performs automatic parameter selection by eliminating predictors with less importance. In high-dimensional data sets, reducing

the set of predictors often leads to better prediction performance and generalizability and has shown improvements over manual stepwise selection methods. This lasso logistic regression model is suited to our platform because it is inherently more intuitive than previous methods which consider hundreds of measurements from cytology that are difficult to interpret.

A lasso logistic regression approach was used to prevent overfitting, reduce coefficient variability, and retain a sparse model with improved generalizability and interpretability. Subjects were dichotomized into “case” and “non-case” outcomes according to their lesion determination (non-case for benign lesions and case for [mild, moderate, severe] dysplasia and malignant lesions). Only subjects with evaluable data for all biomarker measurements and PMOL status were considered (N = 365). Algorithm results were recorded for 241 benign lesion and 124 dysplasia and malignant lesion subjects. Diagnostic performance was characterized by area under the curve (AUC), sensitivity, and specificity. The results from six molecular biomarker assays on the POCOCT system were pooled to obtain final estimates. A receiver operating characteristic (ROC) curve was plotted for the cross-validated test set. Non-zero lasso logistic regression coefficients were retained for the following predictors: percentage of non-mature squamous cells, percentage of small round cells, percentage of leukocytes, age, sex, smoking pack years, lesion major axis diameter, clinical impression of lichen planus, and lesion color (red, white, or red/white) (see **Table S3**). Boxplots of cross-validated algorithm results were generated for the test set responses for benign, mild dysplasia, moderate/severe dysplasia, and malignant lesions. Median numerical indices were compared for each diagnostic classification using a two-sided Wilcoxon rank sum test at a significance level of $p = 0.05$. Internal calibration was performed by sorting and grouping the predicted responses (i.e., numerical index) into deciles and measuring the observed proportions of dysplasia/malignant lesions in each decile. The Hosmer-Lemeshow goodness of fit statistic was used to assess the model fit.

Following this same method, diagnostic algorithms for mild versus moderate dysplasia (3|4), low versus high risk (4|4), moderate versus severe dysplasia (4|5), healthy control (no

lesion) versus malignant (0|6), and benign dysplasia versus malignant (2|6) were also developed. Model responses for each subject were averaged over all biomarker assays to inform diagnostic performance. AUC, sensitivity, and specificity were reported as mean and 95% confidence interval values for the cross-validated test set.

Table S3. Predictor definitions

Abbreviation	Reference	Details
1-%TYPE 1	percentage of non-mature squamous cells	1 – (number of mature squamous cells / total cells), where 'total cells' is the number of cells Types 1-3
%TYPE 2	percentage of small round cells	number of small round cells / total cells, where 'total cells' is the number of cells Types 1-3
%TYPE 3	percentage of leukocytes	number of leukocytes / total cells, where 'total cells' is the number of cells Types 1-3
AGE	age	age in years
SEX	sex	male = 1, female = 0
PACKYR	calculated pack years	average cigarettes smoked per day times years smoked divided by 20
LSIZEMAX	lesion size in maximum dimension	lesion diameter along the long axis in mm
LICHENFN	clinical impression of lichen planus	binary measure completed by clinician at time of brush cytology sample collection indicating the presence ("1") or absence ("0") of the clinical features of lichen planus
LESIONCOLOR	lesion color (red, white, or red/white)	variable indicating lesion color; white = 0, red = 1, red and white = 2

Supplemental Methods References

1. Vigneswaran N, Beckers S, Waigel S, et al. Increased EMMPRIN (CD 147) expression during oral carcinogenesis. *Exp Mol Pathol.* 2006;80(2):147-159.

2. Torres-Rendon A, Roy S, Craig GT, Speight PM. Expression of Mcm2, geminin and Ki67 in normal oral mucosa, oral epithelial dysplasias and their corresponding squamous-cell

carcinomas. *Br J Cancer.* 2009;100(7):1128-1134.

3. Weigum SE, Floriano PN, Redding SW, et al. Nano-bio-chip sensor platform for examination of oral exfoliative cytology. *Cancer Prev Res.* 2010;3(4):518-528.

4. Li HX, Zheng JH, Fan HX, Li HP, Gao ZX, Chen D. Expression of alphavbeta6 integrin and collagen fibre in oral squamous cell carcinoma: association with clinical outcomes and prognostic implications. *J Oral Pathol Med.* 2013;42(7):547-556.

5. Ylipalosaari M, Thomas GJ, Nystrom M, et al. Alpha v beta 6 integrin down-regulates the MMP-13 expression in oral squamous cell carcinoma cells. *Exp Cell Res.* 2005;309(2):273-283.

6. Yu YH, Morales J, Feng L, Lee JJ, El-Naggar AK, Vigneswaran N. CD147 and Ki-67 overexpression confers poor prognosis in squamous cell carcinoma of oral tongue: a tissue microarray study. *Oral Surg Oral Med Oral Pathol Oral Radiol.* 2015;119(5):553-565.

7. Daniel FI, Fava M, Hoffmann RR, Campos MM, Yurgel LS. Main molecular markers of oral squamous cell carcinoma. *Applied Cancer Research.* 2010;30(3):279-288.

8. Williams GH, Romanowski P, Morris L, et al. Improved cervical smear assessment using antibodies against proteins that regulate DNA replication. *Proc Natl Acad Sci.* 1998;95:14932-14937.

9. Scott IS, Odell E, Chatrath P, et al. A minimally invasive immunocytochemical approach to early detection of oral squamous cell carcinoma and dysplasia. *Br J Cancer.* 2006;94(8):1170-1175.

10. Speight PM, Abram TJ, Floriano PN, et al. Interobserver agreement in dysplasia grading: toward an enhanced gold standard for clinical pathology trials. *Oral Surg Oral Med Oral Pathol Oral Radiol.* 2015;120(4):474-482.

11. Abram TJ, Floriano PN, Christodoulides N, et al. 'Cytology-on-a-chip' based sensors for monitoring of potentially malignant oral lesions. *Oral Oncol.* 2016;60:103-111.

12. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods.* 2012;9(7):671-675.

13. Carpenter AE, Jones TR, Lamprecht MR, et al. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* 2006;7(10):R100.