Computer-assisted image analysis of the tumor microenvironment on an oral tongue squamous cell carcinoma tissue microarray

Sangjune Laurence Lee a, Michael Cabanero b,c, Martin Hyrcka b,c, Marcus Butler c,d, Fei-Fei Lia,c, Aaron Hansen c,d, Shao Hui Huang a,d, Ming-Sound Tsao b,c, Yuyao Song c,e, Lin Lu c,e, Wei Xuc,e, Douglas B. Chepeha c,f, David P. Goldsteinc,f, Ilan Weinrebc,c, Scott V. Bratmana,c,g, *Corresponding author at: 610 University Ave, Toronto, ON M5G 2M9, Canada. E-mail address: Scott.Bratman@rmp.uhn.ca (S.V. Bratman).

1. Introduction

Oral cavity cancer is estimated to cause 128,000 deaths annually worldwide. Despite advances in diagnostic and therapeutic approaches, mortality rates of oral cavity squamous cell carcinoma (OTSCC), which make up 90% of all oral neoplasms, have not improved significantly during the last 30 years with a 5-year survival of 40–50% [1–3]. The effect of treatment regimen or other prognosis-related factors is often uncertain and controversial in this disease [4].

Prognostic and predictive biomarkers hold the promise to enable more personalized treatment of OTSCC in order to improve cure rates and minimize side effects. In addition to the tumor, node, and metastases (TNM) staging system, recent studies suggest that the tumor immune microenvironment may provide independent prognostic information [5]. OTSCC is among the most highly immune-infiltrated cancer types [6]. The density of tumor infiltrat-
congregating lymphocytes (TILs) and their exact location within the epithelium or stroma within tumors may be prognostic [7]. Moreover, the presence and distribution of TILs and TIL subsets within tumors may be predictive of response to immune checkpoint inhibitors such as antibodies targeting the PD-1/PD-L1 axis [8].

Immunohistochemistry (IHC) is a semiquantitative diagnostic technique that is routinely used in clinical surgical pathology to evaluate various tumor markers including components of the tumor immune microenvironment [9]. Multiple tumor samples may be stained by IHC and evaluated in a consistent manner using tissue microarrays (TMA). Although TMAs can facilitate throughput of tissue staining, manual reading of IHC studies evaluating protein expression on TMAs is labor intensive and prone to bias [10,11]. Biomarker studies on the tumor immune microenvironment have shown conflicting results at times, which could be attributed to the absence of validated and standardized quantification methods [11,12].

The use of computer-assisted image analysis may enable high-throughput quantitative analysis of IHC on TMAs. Studies have demonstrated the use of computer-assisted image analysis in the quantification of cell count [13,14] and automated segmentation of tissue compartments [15]. In order to localize immune cells to tissue compartments, software packages have used digital pattern recognition [16].

We hypothesized that computer-assisted image analysis of IHC can be used for robust quantification and localization of TILs, TIL subsets, and related biomarker-expressing cells within the tumor microenvironment. Our objective was to validate the use of computer image analysis tools using a TMA of OTSCC. The prognostic implications of these immune markers were then explored.

2. Materials and methods

2.1. Study patients

Patients with OTSCC treated at our institution were included in this retrospective cohort study. Each patient had a biopsy sample taken from the primary tumor site in the oral cavity. Patients were treated with definitive surgical resection of primary 2 neck dissection. Decision of adjuvant treatment was determined in multidisciplinary setting and adjuvant radiotherapy with or without concurrent cisplatin chemotherapy was typically given for patient with adverse pathologic features, according to NCCN guidelines such as pT3-4, close resection margins less than 5 mm, lymphovascular invasion (LVI), perineural invasion (PNI), multiple positive lymph nodes, pathological extranodal extension and positive resection margins. The latter two were considered high-risk features and merit addition of chemotherapy [17]. An experienced head and neck pathologist (I.W.) reviewed all cases for diagnosis and grade. The WHO 2005 classification was used, which is also consistent with the WHO 2017 classification. Demographic, clinical, and outcome data were prospectively collected at the point-of-care.

2.2. TMA preparation and Immunohistochemistry

A tissue microarray (TMA) was built from 182 samples of OTSCC, 2 samples each from 91 patients managed at a Canadian tertiary academic oncology center who consented to being included in a clinical database of patients with newly diagnosed previously untreated OTSCC and in whom adequate tissue was available. Specimens were retrieved from the institution's biobank repository and, if not available in the biobank, from archived pathology specimens. All tumor specimens were preserved in formalin fixed blocks. For each tumor, two duplicate 0.6 mm cores of tumor were included in the TMA.

The immune markers shown in Table 1 were used to label the TMA on serial 4 μm thick tissue sections with brown chromogen stain and hematoxylin counterstain. IHC was performed at Princess Margaret Cancer Center Advanced Molecular Profiling Lab (AMPL) core facility. TMA paraffin sections at 4 μm thickness were dried at 60 °C oven overnight before staining. The IHC was performed according to the manufacture's guidelines using BenchMark XT-an automated slide strainer (Ventana Medical System) with standard antigen retrieval (C1, Tris/Borate/EDTA pH8.0, #950-124) or with protease1 (760–2018) for 4 min. The origin and dilution of the antibodies used for IHC were as follows: anti-CD3 rabbit monoclonal (Ventana/Roche, 2GV6) for 30 min; anti-CD4 rabbit monoclonal (Ventana/Roche, SP35), anti-CD8 rabbit monoclonal (Ventana/Roche, SP57), anti-FOXp3 mouse monoclonal (Abcam, 236A/E7) at 1:100 dilution, and anti-IDO mouse monoclonal (Millipore, 1F8.2) at 1:300 dilution for 60 min; and anti-PD-L1 rabbit monoclonal (CST, E1L3N) at 1:100 dilution for 90 min. Ventana Ultraview Universal DAB Detection Kit (#760-500) containing a cocktail of labeled secondary antibodies was then utilized. The complex was visualized with hydrogen peroxide substrate and 3,3’-diaminobenzidine tetrabhydrochloride (DAB) chromogen. The slides were counterstained with Ventana Hematoxylin II and Bluing reagent, dehydrated in graded alcohol, cleared in xylene and coverslipped in Permount. Downstream analysis was performed on those TMA cores that passed initial quality assessment, including adequate sample area (>75% of that expected for a 0.6 mm diameter core) and lack of sectioning artefacts.

2.3. Computer-assisted image analysis

IHC stained TMA slides were digitized with a magnification of 20× using Aperio Scanscope XT (Leica Biosystems Inc., Buffalo Grove, USA). Tissue Studio (Definiens AG, Munich, Germany) was used to enumerate the number of marker-positive (CD3+, CD4+, CD8+, or FOXp3+) cells and to quantify the staining intensity for IDO or PD-L1. Stromal or epithelial (tumor) compartments of each core was identified using the cytokeratin section as a tissue mask. Cell populations were localized to either stroma or epithelium after rigidly registering each core with the stain of interest to the cytokeratin section through rotations and translations using a custom Matlab script as shown in Fig. 1.

| Table 1 | Immunohistochemistry stains used in this study | |
| --- | --- | --- | --- | --- |
| Marker | Description | Antibody clone name | Supplier | Clone | Dilution |
| CD3 | T-lymphocytes | anti-CD3 rabbit monoclonal | Ventana/Roche | 2GV6 | |
| CD4 | helper T-cells | anti-CD4 rabbit monoclonal | Ventana/Roche | SP35 | |
| CD8 | cytotoxic T-cells | anti-CD8 rabbit monoclonal | Ventana/Roche | SP57 | |
| FOXp3 | regulatory T-cells | anti-FOXp3 mouse monoclonal | Abcam | 236A/E7 | 1:100 |
| IDO | Immune suppressive molecule present on dendritic cells, monocytes and macrophages | anti-IDO mouse monoclonal | Millipore | 1F8.2 | 1:300 dilution for 60 min |
| PD-L1 | Ligand for PD-1, immune suppressive molecule | anti-PD-L1 rabbit monoclonal | CST | E1L3N | 1:100 dilution for 90 min |
| AE1/AE3 | Cytokeratin in epithelial tissue | | | | |
2.4. Manual segmentation and IHC scoring

Manual counts of stained cells were performed by a board-certified pathologist (M.C.). The total numbers of CD3+, CD4+, CD8+, and FOXP3+ cells in the epithelial (tumor) and stromal compartments of each TMA core were enumerated. For PD-L1, the intensity and percent of membranous staining of PD-L1+ tumor cells and membranous and/or cytoplasmic staining of immune cells in both epithelial and stromal compartments were used to calculate an H-score for each TMA core. The H-score is calculated as the sum of percent of cells stained at low intensity times 1, percent of cells stained at intermediate intensity times 2, and percent of cells stained at high intensity times 3, with a range of score from 0 to 300 [18]. Clinical PD-L1 scoring methods such as tumour percentage score (TPS), combined positive score (CPS), or immune cell (IC) score were not used as the H-score provided a wider range of scores to accurately describe different staining intensities [19,20]. Nine representative samples (10% of the population), varying in degree of complexity in differentiating epithelium from stroma, were manually segmented to evaluate the robustness of biomarker quantification and localization. Automatic and manual segmentation of the epithelial (tumor) and stroma compartments was performed on each IHC stained section (Fig. 2).

2.5. Statistical analysis

Descriptive statistics are provided with median and range for continuous variables and frequencies and percentages for categorical variables. Demographics and clinical characteristics are compared by Wilcoxon rank sum test for continuous variables and Fisher exact tests for categorical variables. Tissue segmentations were compared between manual and automatic methods using a Dice coefficient measure. The manual FOXP3+ cell counts and PD-L1 H-scores were compared to the automated cell count and H-score using the Pearson correlation coefficient.

Univariate and multivariable Cox proportional hazards models were used to estimate impact of immune markers on OS and DFS. OS and DFS between cohorts were compared by log-rank test. All tests were two-sided. Results were considered significant if the p-value was ≤0.05. All statistical analyses are performed using SAS 9.4 and R (R Foundation, Vienna, Austria).

3. Results

3.1. Patient cohort for TMA image analysis

91 patients with OTS CC underwent treatment from 2005 to 2008 consisting of surgery only (n = 59) or followed by adjuvant radiotherapy or cisplatin-based chemoradiotherapy (n = 32) (Table 2). 37 patients were never smokers. A TMA containing 2 cores per tumor was stained with 7 different antibodies by IHC for image analysis (total of 1274 tumor images). 241 (18.9%) of the images were excluded due to inadequate sample area, and 7 (0.5%) samples were excluded due to sectioning artifacts. The remaining 1026 (80.5%) images were used for downstream analysis.

3.2. Segmentation of epithelial and stromal compartments

Segmentation of TMA images into epithelial (tumor) and stromal components is needed for detailed characterization of the tumor microenvironment. We compared our automated computer-assisted tissue segmentation procedure that relied on a separately stained cytokeratin mask to manual analysis performed...
on each TMA section (Fig. 2). Agreement between automated and manual tissue segmentation results were variable (Dice coefficient range 0.45–0.95; mean ± standard deviation 0.70 ± 0.15).

We next explored factors that influenced the accuracy of segmentation. Automated segmentation of the epithelial compartment declined in accuracy with increasing distance from the cytokeratin section and with decreasing epithelial tissue area (Figs. 3 and 4). Agreement between automated and manual segmentation was higher when both were performed on the same cytokeratin section than when the automated segmentation was performed on the cytokeratin section and the manual segmentation was performed on a different section taken from the same core (Dice coefficient mean ± standard deviation 0.85 ± 0.06 vs. 0.70 ± 0.15, p = 0.002).

3.3. Comparison of automated versus manual IHC stain scoring on the TMA

We conducted a comparative analysis of automated versus manual biomarker quantification within the tumor microenvironment. For this, we used FOXP3 and PD-L1 to represent cell count and continual staining intensity (H-score), respectively. FOXP3+ cell count and PD-L1 H-score were determined using automated and manual methods for 149 (81.9%) and 143 (78.6%) of the samples with adequate cores (sufficient area and no sectioning artifacts), respectively. As shown in Fig. 5, automated and manual FOXP3+ cell counts had a correlation of $R = 0.56 \ (p = 2.5 \times 10^{-12})$ in the epithelium and $R = 0.90 \ (p = 5.8 \times 10^{-34})$ in the stroma. Automated and manual PD-L1 H-score had a correlation of $R = 0.46 \ (p = 5.0 \times 10^{-8})$ in the epithelium and $R = 0.51 \ (p = 3.0 \times 10^{-10})$ in the stroma. These results suggest that automated computer-assisted image analysis of this OTSCC TMA produces cell counts and H-score results that correlate with manual quantification.

3.4. Automated computer-assisted image analysis of OTSCC tumor microenvironment

Next, we used the automated computer-assisted image analysis of the OTSCC TMA to evaluate putative biomarkers within the tumor microenvironment. Specifically, we assessed the correlation of the infiltrating immune cell counts and staining intensities within the segmented epithelial and stromal compartments (Fig. 6). In the stromal compartment, CD3+, CD4+, CD8+ and FOXP3+ cell counts were strongly correlated with one another (Pearson $R = 0.6$–0.95; $p < 1.0 \times 10^{-15}$ for all), and PD-L1 H-score and CD8+ cell counts were moderately correlated with one another.
In contrast, in the epithelial compartment, CD3+, CD4+, CD8+ and FOXP3+ cell counts had weak correlations in the epithelium (R < 0.3; p ranging between 1.0 × 10⁻³ to 0.9).

Compared to never-smokers, current and ex-smokers had a significantly increased number of CD3+, CD4+, and FOXP3+ cells in the epithelial compartment (OR = 1.003, 1.004, 1.023 p = 0.036, 0.017, 0.022 respectively), and had a trend towards stronger PD-L1 staining in both epithelial and stromal compartments (OR = 1.025, 1.019; p = 0.079, 0.053 respectively); these results were confirmed by manual scores (Supplementary Tables S1–S2). No significant correlations were found between the immune markers of interest and overall stage, N stage, and T stage.

3.5. Prognostic significance of TILs in OTSCC

We next performed an exploratory analysis to identify prognostic biomarkers in this OTSCC cohort. Few statistically significant correlations were observed (Tables S3–S4). Among patients treated with adjuvant radiotherapy (n = 27), high levels of infiltrating (i.e., in the epithelial compartment) CD8+ cells was associated with inferior OS (HR = 1.003, 95% CI: 1.000–1.005, p = 0.03) and DFS (HR = 1.002, 95% CI: 1.000–1.005, p = 0.04) (Fig. 7 and Supplementary Figs. S1–S2). There was a statistically significant interaction effect between epithelial CD8+ cell count and treatment with adjuvant radiotherapy (p = 0.016 for OS, p = 0.0025 for DFS) (Table S5).

4. Discussion

Understanding the distribution and location of immune cells and immune system signaling cell markers within the microenvironment of head and neck cancers may help identify prognostic and predictive biomarkers. Manual scoring of IHC results is labor intensive and susceptible to interobserver and intraobserver variability. Standardization of IHC immune marker scoring through automation for high-throughput analysis of TMAs may facilitate the discovery of these biomarkers. In this work, we demonstrate the use of computer-assisted image analysis on a 91-patient OTSCC TMA to automatically segment epithelial and stromal tissue compartments and determine immune cell count and cell staining intensity. From our exploratory analysis, we show that infiltrating immune cell density and PD-L1 staining intensity is higher in smokers compared to non-smokers. Furthermore, we show that among patients treated with adjuvant radiotherapy, high levels of infiltrating CD8+ cells were associated with a detriment in OS and DFS.

Various studies have demonstrated the use of computer-assisted image analysis to quantify tissue segmentation [16], marker staining intensity [15,21,22] and cell count [13,14,23,24]. Smaller studies have shown image analysis algorithms are more accurate than manual IHC scoring [25]. Despite the abundance of these studies as well as open source and commercially available software, including some FDA approved software packages, the College of American Pathologists and American Society of Clinical Oncology has stated that there is no universally acceptable procedure for validating digital imaging methods [26].

We used Definiens image analysis software as it has been reported to accurately segment tissue compartments and has been used in colorectal cancer “Immunoscore” analysis [25,27,28]. However, in our experience, Definiens digital pattern recognition was unable to consistently identify epithelial and stromal compartments of OTSCC from IHC slides. To improve performance, in this study we utilized a mask from a separate tissue section stained (R > 0.6; p < 1.0 × 10⁻³).
with pan-cytokeratin antibodies in order to distinguish the epithelial tumor compartment from the surrounding stroma. We found that segmentation accuracy was generally good using this approach, but accuracy declined for sections that were serially further from the cytokeratin-stained section or that consisted of a low proportion of the tissue compartment of interest. In future studies, cytokeratin stained sections taken at regular intervals along the length of the core sample could compensate for the expected tissue changes and improve the cytokeratin mask.

Having validated the automated computer-assisted method for tissue compartment segmentation, we evaluated the method for quantification of immune cell counts and biomarker staining intensities within the OTSCC tumor microenvironment. We found that the automated method produced cell counts that were more highly correlated with manual cell counts in the stromal compartment than the epithelial compartment. This may be due to more accurate manual cell counts in the stroma of oral cavity squamous cell carcinomas, where greater numbers of lymphocytes are present as compared with the epithelial compartment [29].

While our methods have yet to be validated with other OTSCC cohorts, our exploratory analysis of the immune IHC scores with clinical factors and outcomes produced interesting findings. There were strong correlations between the stromal populations of CD3+, CD4+, CD8+ and FOXP3+ lymphocytes. In smokers, there were

---

**Fig. 5.** Automated scoring versus manual scoring of cell counts and staining intensities. (a) FOXP3+ cell count in the epithelial, (b) FOXP3+ cell count in the stromal compartment, (c) PD-L1 H-score in the epithelial compartment, and (d) PD-L1 H-score in the stromal compartment.

**Fig. 6.** Relationships between TIL subsets and biomarker expression. Pearson correlation coefficient comparing cell counts for IDO+, CD3+, CD8+, FOXP3+ cells and H-Score for PD-L1 in (a) the epithelial compartment and (b) the stromal compartment.
higher levels of CD3+, CD4+, CD8+ and FOXP3+ lymphocytes and expression of PD-L1. These findings may reflect a higher level of inflammation and immune tolerance in these patients. Lin et al have shown a correlation with high PD-L1 expression and metastasis and poor prognosis in OTSCC [30]. However the effect of smoking on overall prognosis in oral cavity squamous cell carcinoma is controversial [31,32].

Interestingly, we found a significant interaction between the number of infiltrating CD8+ cells and adjuvant radiotherapy on both OS and DFS. High numbers of infiltrating CD8+ TILs portend a worse prognosis specifically for patients treated with adjuvant radiotherapy in this cohort. While previous studies have associated a high CD8+ count with a better prognosis, few studies have focused on the spatial distribution of CD8+ (and other immune cells) in OTSCC with large patient sample sizes [33–37]. Some studies have shown increased tumor recurrence in patients with higher infiltrating CD8+ cells [38,39], while others have shown that CD8+ count does not impact prognosis and that the ratio of CD8+ cells to regulatory immune cells is more prognostic instead [40,41]. There is evidence that radiation alters the ratios of various immune cells [29,42]. We speculate that while abundant TILs counteract tumor progression, radiotherapy may alter the tumor microenvironment so that the tumor could evade detection by the immune system and avoid elimination. This hypothesis generating result will require additional validation studies and mechanistic analysis.

Further studies of the immune microenvironment of OTSCCs will be needed to determine the prognostic and predictive value and clinical utility of potential biomarkers. These future studies may employ a version of the computer-assisted image analysis that we have evaluated in this study. In order to produce robust results, important findings from our study should be taken into consideration, such as the need for close proximity of pan-cytokeratin-stained sections to create a mask for the purpose of automated tissue compartment segmentation. To circumvent this issue, future studies may also employ immunofluorescence imaging and chromogenic multiplexed immunohistochemistry [43], which could produce improved localization of immune markers. In addition, obtaining samples from both the tumor core and the tumor invasive margin may provide further prognostic information [44,45].

Declaration of Competing Interest

Each author confirms there are no business relationships that might lead to a conflict of interest nor any other competing interests.

Acknowledgments

This study was supported by a Strategic Training in Transdisciplinary Radiation Science for the 21st Century (STARS21) training program fellowship awarded to SLL. SVB is supported by the Gattuso-Slaight Personalized Cancer Medicine Fund at the Princess Margaret Cancer Centre. We acknowledge the support from the Princess Margaret Cancer Foundation and the Princess Margaret Cancer Center Head & Neck Translational Program, with philanthropic funds from the Bartley-Smith/Wharton Family, The Joe & Cara Finley Center for Head & Neck Cancer Research, Dr. Mariano Elia, and Gordon Tozer.

Disclosure of interest

Authors have no disclosures.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ctro.2019.05.001.
