ORIGINAL PAPER

A 10-Year Review of Intraoral Salivary Gland Tumor Diagnoses: Diagnostic Challenges and Inter-Observer Agreement

Jessie Fuoco[1] · Mei Dong[2] · Christina MacMillan[3] · Ipshita Kak[4] · Bayardo Perez-Ordonez[5] · Grace Bradley[1,6] · Wei Xu[2] · Marco Magalhaes[1,6]

Abstract

Background Salivary gland tumors (SGT) are a diverse group of neoplasms arising from the major and minor glands. The oral cavity is the most common site for minor SGT (IMSGT), and these lesions frequently pose a challenge to the pathologist due to overlapping histopathological features and limited material for analysis. Our objective was to determine specific clinical and histopathological features associated with challenges in IMSGT diagnoses and pathologists' agreement.

Methods We conducted a retrospective analysis of 248 IMSGT received between 2010 and 2019. We evaluated the diagnostic challenge of the cases by stratifying according to whether a definitive, favored, or indeterminate (challenging) diagnosis was provided. Inter-observer agreement and concordance of biopsy diagnoses with the final diagnoses after tumor resection were evaluated.

Results Of the 248 biopsies, 191 had a definitive diagnosis, 38 favored diagnoses, and 19 were indeterminate. The predominant diagnoses considered for the indeterminate category were pleomorphic adenoma/myoepithelioma (PA), polymorphous adenocarcinoma (PAC), adenoid cystic carcinoma (AdCC), and low-grade adenocarcinoma. Using multivariate analysis of clinical features, younger patient age, smaller tumor size, and larger biopsy size increased the likelihood of a definitive diagnosis ($p = 0.014$, $p = 0.037$, $p = 0.012$). The inter-observer agreement for 68 representative cases was moderate overall (Fleiss's Kappa 0.575) and good for the 40 cases with a definitive diagnosis (Fleiss's Kappa 0.66). Sixty-five biopsy diagnoses were matched with corresponding tumor resection diagnoses and found to show a good concordance (Cramer's V test 0.76). The discordant diagnoses predominantly involved PA, carcinoma exPA, PAC, AdCC, and adenocarcinoma NOS.

Conclusion Diagnostic challenges in IMSGT incisional biopsies were infrequent, especially if multiple pathologists were consulted. PA, PAC, AdCC, and adenocarcinoma NOS were the histologic types more commonly posing diagnostic challenges. Younger patient age, smaller tumor size, and larger biopsy are associated with a definitive diagnosis. This data highlights the importance of appropriate sampling in IMSGT.

Keywords Intraoral · Salivary · Neoplasm · Tumor · Malignancy · Pathology · Concordance · Inter-observer

Introduction

Salivary gland tumors (SGT) represent a diverse group of diseases that differ in clinical behavior, and correct diagnosis is essential for management and prognosis [1, 2]. The most recent WHO classification of SGT includes 15 benign

✉ Marco Magalhaes
  Marco.Magalhaes@dentistry.utoronto.ca

1 Oral & Maxillofacial Pathology & Oral Medicine, University of Toronto, 457-124 Edward Street, M5G 1G6 Toronto, ON, Canada

2 Department of Biostatistics, University Health Network, Princess Margaret Hospital, Toronto, ON, Canada

3 Department of Laboratory Medicine & Pathobiology, Mount Sinai Hospital, Toronto, ON, Canada

4 Department of Laboratory Medicine & Pathobiology, St. Joseph's Healthcare Hamilton, Hamilton, ON, Canada

5 Department of Laboratory Medicine & Pathobiology, University Health Network, Toronto, ON, Canada

6 Department of Dental and Maxillofacial Sciences, Sunnybrook Health Sciences Centre, Toronto, ON, Canada

and 21 malignant epithelial tumors [3]. Overall, 20% of all SGT arise in minor salivary glands, and the oral cavity is the most common location for minor salivary gland tumors [4, 5]. Intraoral minor salivary gland tumors (IMSGT) typically present as slowly enlarging, painless mucosal swellings. It is difficult to differentiate benign from malignant tumors based on clinical presentation, particularly for early lesions [2]. The initial investigation of IMSGT requires an incisional biopsy except for small, encapsulated lesions in the lip or anterior buccal mucosa [6]. The distribution of histologic types varies between major and minor SGT, and there is a higher proportion of malignant tumors in minor glands [2, 7]. Fifty percent of SGT are malignant in minor glands, compared to approximately 25% in the parotid glands [8, 9]. Some tumors are primarily seen in minor salivary glands, including canalicular adenoma, polymorphous adenocarcinoma (PAC), cribriform adenocarcinoma of the salivary glands (CASG), and clear cell carcinoma (CCC). Others preferentially occur in the major salivary glands, including basal cell adenoma, basal cell adenocarcinoma, acinic cell carcinoma (AcCC), salivary duct carcinoma (SDC), intraductal carcinoma, and epithelial-myoepithelial carcinoma [2, 7].

The diagnosis of IMSGT relies primarily on histomorphologic examination, and the presence of overlapping histopathological features between tumors can create diagnostic challenges [10]. Many different SGT share light microscopic and immunohistochemical features, showing overlapping growth patterns, cytomorphologic features, and stromal characteristics. Malignant SGT may show bland cytologic features with few mitoses, so the distinction between benign and malignant is dependent on the identification of invasive growth. This is often a challenge in incisional biopsies due to the limited tissue available and many potential pitfalls [2, 7, 11]. Most IMSGT are unencapsulated, limiting the ability to assess invasion, and conversely, both benign and malignant tumors may appear well circumscribed [2].

Furthermore, the extension of the tumor to the surface epithelium may not represent invasive growth, and certain tumors such as pleomorphic adenoma (PA) can show fatty stromal metaplasia, which, if not recognized, can be easily mistaken for invasion of the tumor cells into fatty tissue. Small intraoral biopsies can suffer from tissue distortion/fragmentation, leaving little viable material to evaluate and limited material for ancillary testing. Although genetic analyses have become increasingly available with the discovery of critical molecular and genomic changes in SGT, high costs are associated with these tests that limit their widespread application [7].

These challenges of IMSGT diagnoses may lead to disagreement among pathologists during the assessment of these tumors and affect patient care. There have been very few studies of inter-observer agreement in IMSGT diagnosis. A recent inter-observer study by Xu and colleagues investigated the level of agreement among 25 expert head and neck pathologists on the specific problem of classifying PAC vs. CASG. Despite the ongoing debate on whether CASG represents a separate entity or variant of PAC, the results showed a fair inter-observer agreement in classifying these tumors (κ=0.370) [12]. Many studies have evaluated the concordance between fine-needle aspiration cytology of salivary gland lesions and the histological diagnosis [13]. Still, virtually no studies evaluated the concordance between the incisional biopsy diagnosis and the surgical pathology (resection) diagnosis of SGT.

We hypothesize that specific clinical and histopathological features of IMSGT are associated with diagnostic challenges in incisional biopsies. To test our hypothesis, we conducted a retrospective analysis of 248 IMSGT biopsies received at an outpatient oral pathology service from 2010 to 2019 inclusive. We stratified cases according to whether a definitive, favored, or indeterminate diagnosis was provided and evaluated clinical and histopathological parameters to determine which factors were associated with an increased likelihood of achieving a definitive diagnosis and which factors were associated with a higher risk of an indeterminate diagnosis. We also evaluated the inter-observer agreement on IMSGT biopsies and concordance of biopsies with the final diagnoses after tumor resection.

Materials and Methods

Design This is a retrospective study of IMSGT biopsies from the Toronto Oral Pathology Service (TOPS). Post-biopsy data were collected from the submitting clinicians and the Pathology departments of St. Joseph's Healthcare Hamilton (SJHH), Hamilton Health Sciences Centre (HHSC), Mount Sinai Hospital (MSH), Sunnybrook Health Sciences Centre (SHSC), and University Health Network (UHN).

Case Selection: We retrieved all cases of IMSGT biopsies received at the Toronto Oral Pathology Service (TOPS) from 2010 to 2019 inclusive. TOPS is an outpatient biopsy diagnostic service within the Faculty of Dentistry, the University of Toronto, that receives biopsies from dentists and dental specialists. Search criteria included terminology from the WHO classification of tumors of salivary glands [3] in the "diagnosis" field of the database. Search terms included a combination of character strings (e.g. "myoepi"), partial diagnoses (e.g. "adenocarcinoma") and full diagnostic terms (e.g. adenoid cystic carcinoma). Search terms included: 'sialadenoma', 'ductal papilloma', 'canalicular

adenoma', 'basal cell adenoma', 'pleomorphic adenoma', 'cystadenoma', 'oncocytoma', 'myoepi*', 'adenocarcinoma', 'mucoepidermoid', 'muco-epidermoid', 'acinic cell', 'adenoid cystic', 'secretory', 'clear cell', 'intraductal', and 'carcinosarcoma'. The following search terms did not yield any results: 'oncocytosis', 'warthin', 'monomorphic', 'salivary prolif*', 'sebaceous adenoma', 'sebaceous lymphadenoma', 'primary squamous', 'salivary duct carcinoma', 'small cell', 'lymphoepithelial carcinoma', 'oncocytic carcinoma', 'sialoblastoma', 'sebaceous adenocarcinoma', 'sebaceous lymphadenocarcinoma', and 'cystadenocarcinoma'. Additionally, the following keywords were searched to account for any minor salivary gland tumor for which the diagnosis was not definitive on biopsy: 'salivary neo*', 'salivary tum*', 'salivary carcinoma', 'undifferentiated', 'see description', and 'likely salivary gland origin'. From this initial search inquiry, 313 pathology reports were retrieved, and 28 duplicates were detected and removed. Next, all pathology reports of the selected cases were reviewed to apply the inclusion criteria that specifically limit intraoral minor salivary gland tumor location. Twenty-five cases were removed as extraoral or did not qualify as SGT. Multiple biopsies of the same tumor were identified and consolidated into a single study ID (reducing the total number by 12). This led to our study's final inclusion of 248 IMSGT biopsies.

Data Collection and Classification

Patient demographics (age, gender, ethnicity, smoking status), clinical information (site, clinical tumor size, symptoms, ulceration, duration, gross biopsy size, number of biopsy pieces), and pathology diagnosis were collected. All diagnoses were coded according to the WHO classification of tumors of salivary glands [3], except pleomorphic adenoma and myoepithelioma, which were placed together in one category. An additional category of 'salivary gland tumor, undetermined' (SGT, undetermined) was included when a salivary gland neoplasm was identified, but no specific diagnosis could be rendered on the biopsy, with definitive classification pending the surgical resection specimen. The lesions were classified into 'benign', 'malignant,' or 'undetermined'. All diagnoses were reviewed and categorized according to levels of uncertainty: A single, **definitive diagnosis** was considered the lowest degree of uncertainty (not challenging). A **favored diagnosis**, e.g., "salivary gland tumor, suggestive of" or "favored", indicates cases that posed challenges that prevented total commitment to a single diagnosis and, therefore, a higher degree of uncertainty. An **indeterminate or unclassifiable diagnosis**, e.g. "salivary gland tumor, see the comment," represents the

highest degree of uncertainty and, therefore, a diagnostic challenge. Seven certified oral pathologists signed out all the biopsy cases.

Inter-Observer Agreement Study

For the inter-observer agreement, we focused on tumors that can have overlapping histopathological features, including PAC, adenoid cystic carcinoma (AdCC), and PA. We included all 20 cases of 'polymorphous adenocarcinoma' and 9 cases of 'adenoid cystic carcinoma' from our initial cohort. Twenty of the 99 'pleomorphic adenoma/myoepithelioma' cases were randomly selected by the statisticians (M.D, and W.X.) concerning the diagnostic challenge and gross biopsy size to investigate these variables further. All 19 diagnoses of 'SGT, undetermined' were included as well. This led to a total of 68 cases selected. The three observers (1 head and neck pathologist and 2 oral pathologists) were blinded to the case histories, original diagnoses, and the diagnoses of the other two pathologists. Each observer was separately given the de-identified slides (including deeper sections if they had been ordered) and asked to make a diagnosis from a list that included all WHO SGT diagnoses (the same list of pathology diagnoses used for the retrospective analysis); if they felt this was not possible, they could record their own diagnoses/comments as needed.

Concordance Study

To determine the concordance between the incisional biopsy diagnosis compared with the final surgical pathology diagnosis, our original records were reviewed, and/or we contacted the submitting clinician (REB approvals – U of T, UHN, MSH, HHSC, SJHH, SHSC) for post-incisional biopsy data, particularly the surgical pathology diagnosis for patients who were referred to hospital Head and Neck Surgery service for treatment. Our list of patients was matched with electronic patient charts of four treating institutions where most patients were referred: University Health Network, Mount Sinai Hospital, Sunnybrook Health Sciences Centre, and Hamilton Health Sciences (St. Joseph's Healthcare Hamilton (SJHH)). Patients were identified by name, date of birth, and/or medical record number. For 9 patients, a second incisional biopsy was performed by the same surgeon. For 7 patients, tumor excision was performed by the same surgeon, and the excision specimen submitted to TOPS did not show the remaining tumor. In 50 cases, a pathology review was performed by anatomic pathologists at the treating institution. For 4 patients, a second biopsy was performed by the treating institution. In 87 cases, the

Table 1 Summary of Data Retrieved After Incisional Biopsy

Total number of cases included in study	248
Second incisional biopsy (TOPS)	9 (4%)
Excisional biopsy negative for tumor (TOPS)	7 (3%)
Pathology consult/review by treating institution	50 (20%)
Second incisional biopsy (by treating institution)	4 (2%)
Surgical resection in treating institution	87 (35%)
Negative for tumor	22 (9%)
Final surgical resection diagnosis	65 (26%)

TOPS: Toronto Oral Pathology Service

surgical resection pathology diagnoses were obtained from the treating institutions. In 22 of the 87 cases, the surgical specimen was negative for tumor, leaving 65 cases with a surgical pathology diagnosis (Table 1). There was no follow-up pathology data in the remaining cases because the submitting clinician could not be reached (practice sold, no response, or passed away), or the patient was lost to follow-up. Additionally, benign tumors could be resected in private practice and do not require a referral to a hospital Head and Neck Surgery service. These 65 surgical resection pathology diagnoses were used as the "gold standard" pathology diagnoses; the incisional biopsy diagnoses and inter-observer diagnoses were then compared for concordance studies. All pathology reports (incisional biopsies, pathology reviews, and surgical resections) were reviewed to document whether immunohistochemistry and molecular testing were performed.

Statistical Analysis

Descriptive statistics were used to examine the distribution of IMSGT by patient gender, age at diagnosis, and anatomic location of the tumor. Continuous variables were reported as median and IQR, whereas categorical variables were reported as frequency and percentages. For statistical analysis, age was grouped into decades; several biopsy pieces were grouped into one vs. multiple (fragmented) pieces, and the anatomic location was grouped into palate/fauces vs. others (buccal mucosa, upper lip, retromolar, lower lip, floor of the mouth, intraosseous, and tongue). The Fisher exact test was used to evaluate the distribution of incisional biopsy pathology diagnoses regarding diagnostic challenges. Univariate analyses (Fisher exact test for categorical variables, Wilcoxon rank-sum test for continuous variables, and logistic regression) were run for all 248 samples to identify clinical variables associated with a 'single definitive diagnosis' and 'SGT, undetermined.' Multivariate analyses (logistic regression) were used to determine the association between this group of clinical variables with 'single definitive diagnosis' and with 'SGT, undetermined.' Sensitivity multivariate models were used to assess for possible confounding factors. The Independent-Samples Kruskal-Wallis test was used to

evaluate the distribution of age across the different biopsy pathology diagnoses. We assessed the inter-observer agreement for 68 cases among three observers. Cohen's kappa was used to assess the agreement of pairs of observers, and Fleiss's kappa was applied to assess the overall agreement for the three observers. The kappa statistic ranges from -1 (complete disagreement) to 1 (complete agreement), with a value of 0 indicating agreement expected by chance. Inter-observer agreement was also stratified by a diagnostic challenge as 'single definitive diagnosis given'; the same could not be performed for 'SGT, undetermined' as the number of cases was too small. Similar stratification was done for gross biopsy size < median, vs. gross biopsy size > median and single biopsy piece vs. multiple fragmented biopsy pieces. Landis and Koch's criteria were applied to evaluate the results. Cramer's V test was used to evaluate the concordance between the incisional biopsy pathology diagnosis, the surgical resection pathology diagnosis (gold standard), and concordance between the 3 observers and the surgical resection pathology diagnosis. The Cramer's V varies from 0 (no correlation) to 1 (complete correlation). Univariate analysis was run using logistic regression to analyze predictor variables for the concordant cases versus the discordant cases. For all comparisons, two-sided tests were applied with $p \leq 0.05$ considered statistical significance. All statistical analyses were conducted in R 4.0.3.

Results

Clinical and Histopathological Findings

Table 2 summarizes the patient demographics, anatomic tumor location, and clinical findings. Of the 248 IMSGT biopsies included in our study, 136 (55%) were from females, 107 (43%) were from males, and 5 (2%) were not specified by the submitting clinician. The average age at biopsy was 53 years, with a range of 10 to 92 years old. The most frequent tumor site was the palate/fauces, with 139 cases originating in this location (56%), followed by buccal mucosa (40, 16%) and upper lip (38, 15%). The mean clinical tumor size was 1.5cm (0.3-7cm). The average gross biopsy size (largest piece of aggregate measurement) was 1cm (0.2–4.7cm). Details on smoking status, ethnicity, symptoms, duration, and ulceration were limited and only specified in less than 50% of the cases and excluded from the analysis.

Of the 248 biopsies included in our study, 229 (92%) were classified as a single definitive diagnosis or favored diagnosis, while 19 (8%) had an indeterminate diagnosis (see Material and Methods) (Table 3). Of the 229 pathology diagnoses, a single definitive diagnosis was given for 191 cases

Table 2 Tumor Location, Size and Patient Demographics

Average age (years)	53 (10–92)
Not specified	5 (2%)
Gender	107 (43%)
Male	136 (55%)
Female	5 (2%)
Not specified	
Tumor location	139 (56%)
Palate/fauces	40 (16.1%)
Buccal mucosa	38 (15.3%)
Upper lip	8 (3.2%)
Retromolar	7 (2.8%)
Lower lip	7 (2.8%)
Floor of mouth	7 (2.8%)
Intraosseous	2 (0.8%)
Tongue	
Average clinical tumor size	1.5cm (0.3-7cm)
Not specified	72 (29%)
Average gross biopsy size	1cm (0.2–4.7cm)
Number of biopsy pieces	130 (52%)
Single piece	118 (48%)
Multiple pieces	

Table 3 Incisional Biopsy Diagnoses

Diagnostic challenge	191 (77%)
Single definitive diagnosis	38 (15%)
Favored diagnosis	19 (8%)
SGT, undetermined	
Benign	143 (58%)
Malignant	91 (37%)
Undetermined	14 (6%)
Pathology diagnosis	99 (39.9%)
Pleomorphic adenoma/myoepithelioma	42 (16.9%)
Mucoepidermoid carcinoma	20 (8.1%)
Polymorphous adenocarcinoma	16 (6.5%)
Cystadenoma	16 (6.5%)
Canalicular adenoma	9 (3.6%)
Adenoid cystic carcinoma	6 (2.4%)
Sialadenoma papilliferum	3 (1.2%)
Ductal papilloma	3 (1.2%)
Acinic cell carcinoma	3 (1.2%)
Clear cell carcinoma	3 (1.2%)
Adenocarcinoma, NOS	2 (1.2%)
Basal cell adenoma	2 (0.8%)
Secretory carcinoma	1 (0.4%)
Oncocytoma	1 (0.4%)
Epithelial-myoepithelial carcinoma	1 (0.4%)
Myoepithelial carcinoma	1 (0.4%)
Intraductal carcinoma	1 (0.4%)
Carcinosarcoma	19 (7.7%)
SGT, undetermined	

SGT: salivary gland tumor

Table 4 Diagnostic Considerations for the 'SGT Undetermined' Cases

Pleomorphic adenoma (2 cases)
Basal cell adenoma
Pleomorphic adenoma, polymorphous adenocarcinoma (2 cases)
Pleomorphic adenoma, low grade adenocarcinoma (no diagnostic and/or actionable fusion gene(s) identified)
Adenoid cystic carcinoma and benign salivary gland tumor mimics
Myoepithelioma/pleomorphic adenoma, polymorphous adenocarcinoma, less likely adenoid cystic carcinoma
Epithelial-myoepithelial carcinoma, adenoid cystic carcinoma, adenocarcinoma NOS, and cellular pleomorphic adenoma
Mucoepidermoid carcinoma, low grade (3 cases)
Polymorphous adenocarcinoma, mucoepidermoid carcinoma
Polymorphous adenocarcinoma, adenoid cystic carcinoma (2 cases)
Myoepithelial tumor, possibility of low grade malignancy cannot be excluded
Basaloid salivary gland neoplasms, basal cell adenoma, polymorphous adenocarcinoma
Carcinoma, likely salivary gland origin, clear cell carcinoma of salivary gland
Adenocarcinoma, likely of salivary gland origin, low grade

'SGT, undetermined', 18 different incisional biopsy diagnoses were rendered. Pleomorphic adenoma/myoepithelioma was the most frequent IMSGT, with 99 cases (40%). Mucoepidermoid carcinoma was the second most frequent IMSGT and the most frequent malignant IMSGT, with 42 cases (17%) (Table3). Table4 lists the differential diagnoses considered in the 19 'SGT, undetermined' cases; notably, 8 cases included both benign and malignant tumors in their differential.

Challenges in Histopathological Diagnosis of Incisional Biopsies

The distribution of diagnoses regarding diagnostic challenges is shown in Table5. There was less uncertainty in the diagnosis of mucoepidermoid carcinomas, seen as a significantly higher proportion of cases with a 'single definitive diagnosis' (40/191; 21%) compared to 'favored diagnosis' (2/38; 5%) (Fisher exact test; $p = 0.021$). On the other hand, although the limited number of cases, basal cell adenomas were associated with more uncertainty, with both cases in the 'favored diagnosis' group ($p = 0.027$). Cystadenomas, PAC, and AdCC also showed a higher proportion in the 'favored diagnosis' group than the 'single definitive diagnosis,' but the differences were not statistically significant. The differential diagnoses mentioned in the reports that gave a 'favored diagnosis' are listed in **Suppl. Table1.** Eighteen of the 248 incisional biopsies mentioned internal or external consultations. One case was "indeterminate" (1/18), 10 cases were "favored" (10/18) and seven cases were given a single diagnosis (7/18). This is likely an underestimation of

(77%), and a diagnosis was favored in 38 cases (15%). 143 of the 248 biopsies represented benign (58%) neoplasms, 91 were malignant (37%), and 14 cases were undetermined as to whether the tumor was benign or malignant (6%); this led to a total of 105 potentially malignant cases. From these 105 cases, 87 were identified in the treating institutions and 65 showed tumor in the resection specimen. Excluding

Table 5 Distribution of Incisional Biopsy Diagnoses According to Diagnostic Uncertainty

Covariate: Pathology Diagnosis	Full Sample (n = 229)	Single Definitive Diagnosis (n = 191)	Favored Diagnosis (n = 38)	P-Value*
Ductal Papilloma	3 (1%)	3 (2%)	0 (0%)	1
Cystadenoma	16 (7%)	12 (6%)	4 (11%)	0.31
Canalicular Adenoma	16 (7%)	13 (7%)	3 (8%)	0.73
Sialadenoma Papilliferum	6 (3%)	6 (3%)	0 (0%)	0.59
Pleomorphic Adenoma / Myoepithelioma	99 (43%)	86 (45%)	13 (34%)	0.28
Basal Cell Adenoma	2 (1%)	0 (0%)	2 (5%)	**0.027**
Oncocytoma	1 (0%)	1 (1%)	0 (0%)	1
Mucoepidermoid Carcinoma	42 (18%)	40 (21%)	2 (5%)	**0.021**
Polymorphous Adenocarcinoma	20 (9%)	15 (8%)	5 (13%)	0.34
Adenoid Cystic Carcinoma	9 (4%)	6 (3%)	3 (8%)	0.17
Acinic Cell Carcinoma	3 (1%)	1 (1%)	2 (5%)	0.072
Secretory Carcinoma	2 (1%)	2 (1%)	0 (0%)	1
Clear Cell Carcinoma	3 (1%)	1 (1%)	2 (5%)	0.072
Epithelial Myoepithelial Carcinoma	1 (0%)	1 (1%)	0 (0%)	1
Myoepithelial Carcinoma	1 (0%)	0 (0%)	1 (3%)	0.17
Intraductal Carcinoma of Salivary Gland	1 (0%)	0 (0%)	1 (3%)	0.17
Adenocarcinoma NOS	3 (1%)	3 (2%)	0 (0%)	1
Carcinosarcoma	1 (0%)	1 (1%)	0 (0%)	1

*Fisher Exact Test

consults as internal informal consults are not documented in the final report.

Clinical Factors Associated with Diagnostic Certainty

The clinical parameters were analyzed using univariate and multivariate analysis regarding the likelihood of having a 'single definitive diagnosis' on incisional biopsy (n = 191/248) (Table 6). In univariate analysis, younger patient age (OR of 0.74, 95% CI (0.61, 0.89), p = 0.0019), smaller clinical tumor size (OR of 0.97, 95% CI (0.94, 1.00), p = 0.03), and a single biopsy piece (OR of 0.44, 95% CI (0.24, 0.81), p = 0.0081) were found to be associated with a single definitive diagnosis. In multivariate analysis, only patient age and clinical tumor size were significant (p = 0.014 for age, p = 0.037 for clinical tumor size). In agreement with our clinical experience, sensitivity multivariate analysis

showed that 'clinical tumor size' is a confounding factor to 'number of biopsy pieces'. The gross biopsy size was not statistically significant in the univariate analysis (p = 0.08) but was in the multivariate analysis (p = 0.012): the larger the gross biopsy size, the more likely the chance of having a single diagnosis. Gender (male vs. female) and anatomic location (palate/fauces vs. other locations) were not statistically significant parameters.

Clinical Factors Associated with Diagnostic Uncertainty

Next, the clinical parameters were analyzed using univariate and multivariate analysis regarding the likelihood of the indeterminate diagnosis 'SGT, undetermined' on incisional biopsy (n = 19/248) (Table 7). Increased patient age was associated with an indeterminate diagnosis (OR of 1.49, 95% CI (1.08, 2.05), p = 0.014). The number of biopsy pieces was not statistically significant in the univariate analysis but showed a trend: a biopsy submitted in multiple fragmented pieces has a higher risk of diagnosis of 'SGT, undetermined' (p = 0.066). Gender, anatomic location, and clinical tumor size were not statistically significant parameters. There was no difference in age distribution in the different diagnostic categories (data not shown).

Inter-Observer Agreement

The overall inter-observer agreement among the three observers for the 68 selected cases was moderate: Fleiss's kappa 0.575, 95% CI (0.5, 0.65). Observers 1 and 2 are certified oral and maxillofacial pathologists, and observer 3 is a certified head and neck pathologist. The 68 cases included 40 cases with a single definitive diagnosis, 9 cases with a favored diagnosis, and 19 cases with the diagnosis of 'SGT, undetermined.' We stratified the inter-observer agreement by the level of diagnostic challenge, but only the 'single definitive diagnosis' group was large enough for analysis, showing good agreement among the three observers with Fleiss's kappa of 0.66, 95% CI (0.55, 0.77), demonstrating an overall better agreement in these cases. When the inter-observer agreement was stratified by gross biopsy size below or above the median (N = 35; N = 33), Fleiss's kappa for the former was 0.59, 95% CI (0.5, 0.7), and 0.57, 95% CI (0.43, 0.65) for the latter, demonstrating a moderate agreement regardless of biopsy size. When the inter-observer agreement was stratified by multiple vs. 1 biopsy piece, Fleiss's kappa was 0.51, 95% CI (0.41, 0.61) with fragmented biopsy pieces and 0.63, 95% CI (0.52, 0.74) with one piece of tissue, demonstrating better overall agreement when the biopsy is submitted as one entire piece instead of multiple fragments. Figure 1 shows representative images of

199

Table 6 Association Between Clinical Parameters and Likelihood of Having a Single Diagnosis

Clinical parameters	Study cohort (n = 248)	Single diagnosis (n = 191)	Univariate (Fisher exact test or Wilcoxon rank-sum test) p-value	Univariate logistic regression OR (95% CI)	Multivariate logistic regression OR (95% CI)
Age (decades)	5.3 (1.7)	5.2 (1.7)	**0.0027**	0.74	0.73
Mean (sd)	5.5 (1, 9.2)	5.3 (1, 9.1)		(0.61,0.89)	(0.57,0.94)
Median (min,max)				**p = 0.0019**	**p = 0.014**
Missing (%)	5 (2)	4 (2)			
Gender	136 (55)	100 (52)	0.12	1	
Female (%)[$]	107 (43)	88 (46)		1.67	
Male (%)	5 (2)	3 (2)		(0.89,3.12)	
Missing (%)				p = 0.11	
Anatomic location	139 (56)	102 (53)	0.13	1	
Palate/fauces (%)[$]	109 (44)	89 (47)		1.61	
Others (%)				(0.87,2.98)	
				p = 0.13	
Clinical tumor size (mm)	15.1 (11)	14 (9.9)	**0.032**	0.97 (0.94,1)	0.96 (0.93,1)
Mean (sd)	10	10		**p = 0.03**	**p = 0.037**
Median (min,max)	(2.5,70)	(2.5,50)			
Missing (%)	72 (29)	56 (29)			
Gross biopsy size (cm)	1 (0.5)	1.1 (0.5)	0.088	1.91	4.08
Mean (sd)	1	1		(0.92,3.96)	(1.35,12.28)
Median (min,max)	(0.2,4.7)	(0.2,4.7)		p = 0.08	**p = 0.012**
Number of biopsy pieces	130 (52)	109 (57)	**0.0099**	1	1
1 (%)[$]	118 (48)	82 (43)		0.44 (0.24,0.81)	0.62
Multiple (%)				**p = 0.0081**	(0.29,1.35)
					p = 0.23

[$]Reference parameter, p = global p-value,

Table 7 Association Between Clinical Parameters and Likelihood of 'SGT Undetermined' Diagnosis

| Clinical parameters	Study cohort (n = 248)	SGT undetermined (n = 19)	Univariate (Fisher exact and Wilcoxon rank-sum test) p-value	Univariate OR logistic regression (95% CI)	Multivariate OR logistic regression (95% CI)
Age (decades)	5.3 (1.7)	6.3 (1.6)	**0.01**	1.51 (1.1,2.07)	1.49
Mean (sd)	5.5	6.3		**p = 0.011**	(1.08,2.05)
Median (min,max)	(1,9.2)	(2,9.2)			**p = 0.014**
Missing (%)	5 (2)	0 (0)			
Gender	136 (55)	13 (72)	0.22	1	
Female (%)[$]	107 (43)	5 (28)		0.46 (0.16,	
Male (%)	5 (2)	1 (5)		1.34) p = 0.16	
Missing (%)					
Anatomic location	139 (56)	14 (74)	0.15	1	
Palate/fauces (%)[$]	109 (44)	5 (26)		0.43	
Others (%)				(0.15,1.23)	
				p = 0.12	
Clinical tumor size (mm)	15.1 (11)	17.9	0.072	1.02	
Mean (sd)	10	(8.9)		(0.98,1.07)	
Median (min,max)	(2.5,70)	20 (4,40)		p = 0.3	
Missing (%)	72 (29)	4 (21)			
Gross biopsy size (cm)	1 (0.5)	0.9 (0.4)	0.39	0.51 (0.15,1.7)	
Mean (sd)	1	0.9		p = 0.27	
Median (min,max)	(0.2,4.7)	(0.3,1.6)			
Number of biopsy pieces	130 (52)	6 (32)	0.092	1	1
1 (%)[$]	118 (48)	13 (68)		2.56 (0.94,6.97)	2.42
Multiple (%)				p = 0.066	(0.88,6.7)
					p = 0.088

[$]Reference parameter, p = global p-value

cases originally diagnosed as SGT, undetermined that were also discordant between the 3 observers.

Concordance Between Incisional Biopsy Diagnosis and Surgical Pathology Diagnosis

The concordance was determined between the 65 incisional

Fig. 1 Representative images of discordant cases diagnosed as "SGT, undetermined" in the original biopsy
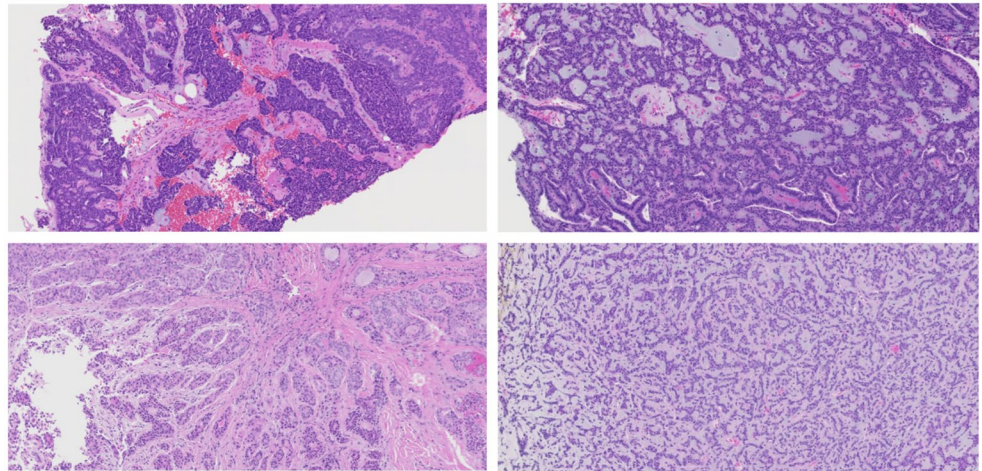
biopsy pathology diagnoses for which their final surgical resection pathology diagnoses could be retrieved. The results showed a good overall concordance of 0.76 (Cramer's V test). Of the 19 'SGT, undetermined' incisional biopsies, 9 surgical resection pathology diagnoses were retrieved; 4 were diagnosed as pleomorphic adenomas/myoepitheliomas, 3 as adenoid cystic carcinomas, 1 as mucoepidermoid carcinoma, and 1 as clear cell carcinoma. Suppl. Table2 collates all the diagnoses obtained for the 19 'SGT, undetermined' incisional biopsies from the inter-observer study, pathology consult or review, second incisional biopsy and surgical pathology. In 5 of 9 cases, a consensus diagnosis agreed with the surgical pathology diagnosis despite the original biopsy diagnosis of 'SGT, undetermined.'

Suppl. Table3 shows discordance between incisional biopsy and surgical resection diagnoses (a total of 56). Seven of 40 (18%) incisional biopsy diagnoses that were 'single definitive diagnosis' were discordant, and 4 of 7 discordances involved PAC. Four of 16 (25%) incisional biopsy diagnoses that were 'favored diagnosis' were discordant, and 2 of 4 discordances involved PAC. One incisional biopsy diagnosed as polymorphous adenocarcinoma was signed out on resection as "low-grade adenocarcinoma," and an incisional biopsy pathology diagnosis favoring pleomorphic adenoma/myoepithelioma was signed out on resection as "spindle cell neoplasm with myxoid and epithelioid features." Of the remaining discordant cases, some were discordant concerning whether the tumor was benign or malignant. Others were discordant with the diagnosis but not whether they were benign or malignant.

IHC and special stains were ordered in 33 of the 248 biopsy cases. 4/33 cases were "indeterminate", 14/33 cases were "favored", and 15/33 were given a "definitive" diagnosis. Molecular testing was performed on 1 biopsy case, confirming a MAML2 rearrangement and allowing for the diagnosis of a low-grade central mucoepidermoid carcinoma. IHC was performed by the treating or consulting

institution in only 25 cases, and molecular testing was performed in 13 cases and were only helpful in the diagnosis of very few cases (e.g. FISH was negative, or there was no diagnostic and/or actionable fusion gene(s) identified).

Univariate analysis, using logistic regression, assessed the clinical and pathological parameters in the 65 cases where concordance could be evaluated. No variable was statistically significant in decreasing the concordance (Suppl. Table4).

Lastly, concordance was also calculated for the inter-observer study, using Cramer's V test. Of the 68 cases studied in the inter-observer agreement, 28 had the surgical resection pathology diagnosis (gold standard). There was good concordance, with a concordance of 0.8, 0.75, and 0.75, for observers 1, 2, and 3, respectively; therefore, the observers' diagnosis was robust.

Discussion

SGT is a diagnostically challenging group of head and neck pathology neoplasms, with a complex classification and overlapping histologic features. IMSGT typically presents as slowly enlarging swellings that are initially investigated by incisional biopsy. The assessment of an incisional biopsy may be hampered by a limited amount of tumor tissue, fragmented specimens that do not allow examination of the tumor borders, and non-representative sampling of a large, heterogeneous tumor. When faced with a problematic biopsy, the pathologist may issue a differential diagnosis, with or without a favored diagnosis, or give an indeterminate diagnosis such as 'Salivary gland neoplasm, undetermined type' [6, 11]. The definitive diagnosis of such cases would require examining the surgical pathology specimen.

We conducted a retrospective review of IMSGT in a 10-year period (2010–2019 inclusive) to study the diagnostic challenges for these tumors. Among 248 cases included

in our study, there were 191 cases with a definitive diagnosis and 38 cases with a 'favored' diagnosis, where a diagnosis was rendered, but one or more other diagnostic possibilities could not be excluded based on the incisional biopsy. The remaining 19 cases had an indeterminate diagnosis such as 'salivary gland tumor, undetermined.' The predominant considerations in the indeterminate cases were pleomorphic adenoma/myoepithelioma (PA), polymorphous adenocarcinoma (PAC), adenoid cystic carcinoma (AdCC), and low-grade adenocarcinoma (Table 4). Comparing SGT distribution in the 'definitive diagnosis' and 'favored diagnosis' groups showed that mucoepidermoid carcinoma (MEC) formed a significantly higher proportion of 'definitive diagnosis' than 'favored diagnosis.' Cystadenoma, PAC, and AdCC showed the opposite pattern with a higher proportion in 'favored diagnosis' than 'definitive diagnosis,' but the differences did not reach statistical significance. Based on the frequent inclusion of PA, PAC, and AdCC in instances of diagnostic challenge, we chose these three entities, along with the indeterminate cases (SGT, undetermined), for our inter-observer agreement study. Adding cases with diagnostic uncertainty (28 cases of 'favored diagnosis' or 'SGT, undetermined') to the 40 cases of 'definitive diagnosis' reduced the level of interobserver agreement from 0.66 (good) to 0.575 (moderate) (Fleiss' kappa). We also compared the incisional biopsy diagnosis and the diagnoses from the inter-observer study with the surgical pathology diagnosis, which is the gold standard diagnosis. Discordance between incisional biopsy diagnosis and surgical pathology diagnosis predominantly involved PAC, followed by PA/carcinoma ex PA, and adenocarcinoma NOS (Suppl. Table 3).

The problem with distinguishing between PA, PAC, and AdCC is well-recognized in SGT diagnosis. These three tumors may all show a mixture of cribriform and tubular growth patterns [2, 6, 14]. The cellular features differ among these tumors but may not be readily appreciated in small and distorted biopsies. These tumors also show different behavior at the periphery since PA is well demarcated from surrounding tissues while PAC and AdCC are invasive. However, incisional biopsies may not include sufficient sampling of the tumor periphery for a confident assessment of invasion. The correct diagnosis by incisional biopsy is needed to guide treatment and prognosis, especially since PA is benign and PAC and AdCC are malignant. PA requires a simple excision, while PAC requires resection that may include adjacent bone. AdCC requires wide resection, and adjuvant radiation therapy is often used when there is concern about positive resection margins [11, 14].

A comparison of the incisional biopsy diagnosis and surgical pathology diagnosis indicated the incisional biopsy diagnosis' accuracy and demonstrated the diagnostic pitfall

in discordant cases. There were 11 discordant cases out of 56 cases (20%) where incisional biopsy and surgical pathology diagnoses could be matched; of which a higher proportion was initially a favored diagnosis (4/16; 25%), compared to a single definitive diagnosis (7/40; 18%) (Suppl. Table 3). As mentioned above, discordance between incisional biopsy diagnosis and surgical pathology diagnosis most often involved PAC, which may be confused with PA, carcinoma ex PA, AdCC, or adenocarcinoma NOS. PAC is a relatively new SGT compared to well-established entities such as PA or AdCC, and its histopathologic and immunohistochemical features have been elucidated in recent years. In particular, immunohistochemical staining for S100, p40, and p63 have become valuable adjuncts to distinguish PAC from PA, AdCC, and low-grade adenocarcinoma [2, 14].

We examined all pathologic diagnoses collected for the 19 incisional biopsies that were 'SGT, undetermined.' For 9 of the 19 cases, the surgical pathology diagnosis was available. Notably, in 4 of 9 indeterminate incisional biopsy cases, there was a consensus from the inter-observer study that agreed with the surgical pathology diagnosis. In one case, the surgical specimen was negative for the tumor. In the remaining 9 cases, surgical pathology diagnosis was not available, but in 5 of 9 cases, there was a pathology consult diagnosis or a second incisional biopsy diagnosis in addition to the diagnoses from the inter-observer study. In 3 of these 5 cases, there was a majority diagnosis (3 of 4 pathologists concurred). These observations suggest that a diagnostic challenge may be resolved through consultation among pathologists. Another consideration is that SGT diagnoses have been refined over the years with better delineation of histomorphologic features for tumors with similar growth patterns and development of diagnostic adjuncts by immunohistochemical staining and molecular genetic studies [2, 7, 11, 15]. For example, the diagnosis of AdCC may be supported by fluorescence in-situ hybridization (FISH) to demonstrate *MYB-NFIB* translocation or immunohistochemical staining for nuclear MYB protein on routine biopsy specimens. However, molecular genetic studies may be limited by costs and availability and do not always lead to a definitive diagnosis [7]. We also considered the experience of the pathologist as a potential factor in reaching a diagnosis of "SGT, undetermined". The 19 cases were signed out by 6 different pathologists and only 2 of the 19 cases were signed out by a less experienced oral pathologist (< 5 years of experience). In addition, our interobserver agreement included an experienced oral pathologist – observer 1 (> 30 years of experience), an oral pathologist with 5 years of experience (observer 2) and an experienced head and neck pathologist (observer 3). The agreement between observers 1 (experienced oral pathologist) and 2 (less experienced oral pathologist) as well as between observer 1 and 3 was identical:

Fleiss's kappa 0.56, 95% CI (0.43, 0.69). The agreement between observer 2 and 3 was good: Fleiss's kapa 0.62, 95% CI (0.48, 0.76) supporting that experience was not a factor in defining a definitive diagnosis.

We studied the effects of clinical parameters of the incisional biopsy on diagnostic challenges. There was a consistent effect of the patient's age in this regard. For every decade decrease in age, there was a higher chance of a definitive diagnosis being given. Conversely, with every decade increase in age, there was a higher likelihood of diagnosis as 'SGT, undetermined.' This could be partly explained by differences in age distribution among different SGT types. In an extensive review of SGT, Wang and colleagues demonstrated an average age of 40 years (range 3–82 years) for patients with a benign tumor and 49 years (range 14–85 years) for patients with a malignant tumor [16]. In our study, the mean age of patients with PA was much lower than that for PAC and AdCC. PA was the most frequent tumor and covered a wide age range from the 1st decade to the 9th decade (data not shown). 86 of 99 cases of PA were in the category of 'single definitive diagnosis,' and 13 were 'favored diagnosis.' We also considered the possibility that increasing age is associated with more significant genetic and phenotypic complexities in tumors, but further investigation of this hypothesis would require studies of age-related changes in SGT biology.

Multivariate analysis of the effect of clinical parameters on the likelihood of a definitive diagnosis showed an association with increasing biopsy size (by gross examination, $p = 0.012$) and decreasing clinical tumor size (by clinical description, $p = 0.037$). A larger incisional biopsy and/or a smaller tumor would logically be more representative of the tumor and more likely to show more of the tumor margin to allow evaluation of tumor invasiveness. There were no clinical parameters besides patient age that were significantly associated with the likelihood of diagnosis as 'SGT, undetermined.' This diagnosis was more likely when the biopsy was in multiple pieces rather than a single piece, but the difference was not statistically significant ($p = 0.066$ by univariate analysis, $p = 0.088$ by multivariate analysis). Interestingly, when the inter-observer agreement was stratified by single versus multiple biopsy pieces, the agreement was improved from moderate for multiple pieces (Fleiss' kappa 0.51) to good for a single piece (Fleiss' kappa 0.63). A biopsy in multiple pieces is likely to be fragmented, making it difficult to assess the growth pattern and invasion at the margins.

Conclusion

In conclusion, diagnostic challenges in IMSGT were relatively infrequent, as identified from the proportion of indeterminate diagnoses and cases with discordance between incisional biopsy and surgical pathology diagnoses. We identified clinical parameters which may help make a definitive diagnosis on incisional biopsy. A more extensive incisional biopsy that is representative of the clinical extent of the tumor and avoids fragmentation of the biopsy will increase the likelihood of a definitive diagnosis.

Data Availability All data is included in the manuscript.

Code Availability Not applicable.

Declarations/Compliance with Ethical Standards

Conflicts of interest/Competing interests The authors declare no conflicts to disclose.

Ethics approval This study was conducted following approval from the Research Ethics Boards of the University of Toronto (Protocol #38153), HHSC and SJHH (HiREB Project #11158), MSH (Protocol #20-0148-C), SHSC (Project #2374), and UHN (CAPCR ID 20-5220).

Consent to participate Not applicable.

Consent for publication Not applicable.

References

1. Moutasim KA, Thomas GJ. Salivary gland tumours: update on molecular diagnostics. Diagn Histopathology. 2020;26(4):159–64. doi:https://doi.org/10.1016/j.mpdhp.2020.01.002.
2. Speight PM, Barrett AW. Salivary gland tumours: diagnostic challenges and an update on the latest WHO classification. Diagn Histopathology. 2020;26(4):147–58. doi:https://doi.org/10.1016/j.mpdhp.2020.01.001.

3. El-Naggar AK, Chan JKC, Rubin Grandis J, Takata T, Slootweg PJ, International Agency for Research on C. WHO classification of head and neck tumours. 2022.

4. Guzzo M, Locati LD, Prott FJ, Gatta G, McGurk M, Licitra L. Major and minor salivary gland tumors. Crit Rev Oncol Hematol. 2010;74(2):134–48. doi:https://doi.org/10.1016/j.critrevonc.2009.10.004.

5. Pires FR, de Almeida OP, Pringle G, Chen SY. Differences on clinicopathological profile from intraoral minor salivary gland tumors around the world. Oral Surg Oral Med Oral Pathol Oral Radiol Endod. vol2. United States2008. p.136–8.

6. Khurram SA, Barrett AW, Speight PM. Diagnostic difficulties in lesions of the minor salivary glands. Diagn Histopathology. 2017;23(6):250–9. doi:https://doi.org/10.1016/j.mpdhp.2017.04.008.

7. Rooper LM. Challenges in Minor Salivary Gland Biopsies: A Practical Approach to Problematic Histologic Patterns. Head Neck Pathol. 2019;13(3):476–84. doi:https://doi.org/10.1007/s12105-019-01010-8.

8. Bishop JA, Thompson LDR, Wakely PE Jr, Weinreb I. Tumors of the Salivary Glands. AFIP Atlases of Tumor and Non-Tumor Pathology. Vol.5. Arlington: The American Registry of Pathology; 2021.

9. Lukšić I, Virag M, Manojlović S, Macan D. Salivary gland tumours: 25 years of experience from a single institution in Croatia. J Craniomaxillofac Surg. 2012;40(3):e75–81. doi:https://doi.org/10.1016/j.jcms.2011.05.002.

10. Todorovic E, Dickson BC, Weinreb I. Salivary Gland Cancer in the Era of Routine Next-Generation Sequencing. Head Neck Pathol. 2020;14(2):311–20. doi:https://doi.org/10.1007/s12105-020-01140-4.

11. Turk AT, Wenig BM. Pitfalls in the biopsy diagnosis of intraoral minor salivary gland neoplasms: diagnostic considerations and recommended approach. Adv Anat Pathol. 2014;21(1):1–11. doi:https://doi.org/10.1097/pap.0000000000000000.

12. Xu B, Barbieri AL, Bishop JA, Chiosea SI, Dogan S, Di Palma S, et al. Histologic Classification and Molecular Signature of Polymorphous Adenocarcinoma (PAC) and Cribriform Adenocarcinoma of Salivary Gland (CASG): An International Interobserver Study. The American Journal of Surgical Pathology. 2020;44(4).

13. Colella G, Cannavale R, Flamminio F, Foschini MP. Fine-needle aspiration cytology of salivary gland lesions: a systematic review. J Oral Maxillofac Surg. 2010;68(9):2146–53. doi:https://doi.org/10.1016/j.joms.2009.09.064.

14. Rooper L, Sharma R, Bishop JA. Polymorphous Low Grade Adenocarcinoma has a Consistent p63+/p40 – Immunophenotype that Helps Distinguish it from Adenoid Cystic Carcinoma and Cellular Pleomorphic Adenoma. Head Neck Pathol. 2015;9(1):79–84. doi:https://doi.org/10.1007/s12105-014-0554-4.

15. Seethala RR, Stenman G. Update from the 4th Edition of the World Health Organization Classification of Head and Neck Tumours: Tumors of the Salivary Gland. Head Neck Pathol. 2017;11(1):55–67. doi:https://doi.org/10.1007/s12105-017-0795-0.

16. Wang D, Li Y, He H, Liu L, Wu L, He Z. Intraoral minor salivary gland tumors in a Chinese population: a retrospective study on 737 cases. Oral Surg Oral Med Oral Pathol Oral Radiol Endod. 2007;104(1):94–100. doi:https://doi.org/10.1016/j.tripleo.2006.07.012.