A Pattern-Based Risk-Stratification Scheme for Salivary Gland Cytology: A Multi-Institutional, Interobserver Variability Study to Determine Applicability

Christopher Griffith, Emory University
Alessandra Schmitt, Emory University
Liron Pantanowitz, University of Pittsburgh
Sara E. Monaco, University of Pittsburgh

Journal Title: Cancer
Volume: Volume 125, Number 10
Publisher: Wiley: 12 months | 2017-10-01, Pages 776-785
Type of Work: Article | Post-print: After Peer Review
Publisher DOI: 10.1002/cncy.21906
Permanent URL: https://pid.emory.edu/ark:/25593/tdq0t

Final published version: http://dx.doi.org/10.1002/cncy.21906

Copyright information:
© 2017 American Cancer Society

Accessed June 4, 2019 1:40 AM EDT
A Pattern-Based Risk Stratification Scheme for Salivary Gland Cytology: A Multi-Institutional Interobserver Variability Study to Determine Applicability

Christopher C. Griffith, MD, PhD, Alessandra C. Schmitt, MD, Liron Pantanowitz, MD, and Sara E. Monaco, MD

1Department of Pathology, Emory University, Atlanta, GA
2Department of Pathology, University of Pittsburgh Medical Center, Pittsburgh, PA

Abstract

Background—Salivary gland aspiration cytology is useful in preoperative management of patients but remains challenging due to extensive morphologic overlap of some tumors limiting the ability to always determine the presence of malignancy. In response to this challenge, there has been increasing drive to develop a risk-based categorization scheme for salivary gland aspirates. Herein we examine the interobserver variability of one such pattern and risk-based system.

Methods—Select smears and cell block sections from 50 salivary gland aspirates from two large academic centers were digitally imaged. These scanned slides were independently and blindly reviewed by four cytopathologists and each aspirate was assigned to one of the proposed pattern-based categories if it was considered neoplastic by the observer. Interobserver agreement was scored and aggregated risks of malignancy calculated for cases with available surgical follow-up.

Results—42 (84%) were considered neoplastic by at least two observers and scored for interobserver agreement. 23.8% (10 of 42) had uniform agreement, 33.3% (14 of 42) majority agreement, and 11.9% (5 of 42) divided agreement. Only 21.4% (9 of 42) had minimal agreement and 9.5% (4 of 42) had no agreement. Condensation of similar categories was able to improve interobserver agreement and still maintain stratified risk of malignancy.

Conclusions—This proposed pattern-based risk stratification scheme, which could be implemented with the forthcoming Milan System, has good overall interobserver agreement and successfully stratifies risk of malignancy. Some simplification is possible to make the system easier to use and improve interobserver agreement while maintaining stratification of risk.
Keywords
Salivary Gland Neoplasms; Biopsy; Fine-Needle; Observer Variation; Risk Stratification

Introduction
Aspiration cytology of salivary gland tumors is useful as a preoperative tool for selecting patients who are in need of surgical excision and in determining the extent of surgery required for such patients. Salivary gland fine needle aspiration (FNA) is a cost effective triage tool and can help reduce the number patients undergoing surgery. In a systematic review and meta-analysis of the diagnostic accuracy of FNA for parotid gland lesions, Schmidt et al estimated 96% sensitivity and 98% specificity in the distinction of neoplastic versus non-neoplastic tumors. The distinction of benign versus malignant in this review article was estimated to have a lower sensitivity of 80% and specificity of 97%. A more recent meta-analysis has shown similar estimates for the ability of parotid FNA to differentiate benign versus malignant (sensitivity 78% and specificity 98%). Although both of these studies show good sensitivity and specificity there is considerable heterogeneity among individual studies ranging from 0% to 100% for sensitivity and 67% to 100% for specificity.

Despite the good performance of salivary gland FNA reported in the literature, many challenges remain in this area. The primary challenge is the difficulty experienced in specifically diagnosing tumor entities, which can be attributed to the extensive morphologic overlap of some salivary gland tumors. This morphologic overlap is perhaps most prominent in the area of basaloid neoplasms and is exemplified in a recent College of American Pathologists (CAP) survey focusing on the cytologic recognition of adenoid cystic carcinoma (AdCC). This study concluded that the accurate identification of AdCC on cytologic samples is a problematic area with a false negative rate for malignancy of 36.4%. Misinterpretation of aspirates from AdCC as pleomorphic adenoma (PA) or monomorphic adenoma was the most common reason for false negative results.

There are similar challenges in the diagnosis of oncocytoid neoplasms and this is best demonstrated in the significant overlap that can be seen in aspirates of low grade mucoepidermoid carcinoma (MEC) and Warthin tumor (WT) with mucinous metaplasia. As the diagnosis of WT should prompt simple excision or even just observation, the misdiagnosis of a MEC as WT on aspiration cytology could result in disease progression in some patients. Similarly, the diagnosis of acinic cell carcinoma (AciCC) is also a potential source of false negative cytology due to the frequent bland appearance of tumor cells which can mimic non-neoplastic salivary acini. The recent description of secretory carcinoma of salivary glands also highlights the morphologic overlap with AciCC.

While studies examining specific morphologic features are available in the literature regarding the distinction of some salivary gland tumors from morphologic mimics, there has been a more recent trend to develop a risk stratification approach to classify salivary gland aspirates to help guide management. In a retrospective multi institutional international study, it was shown that separation of salivary gland aspirates into primary categories of neoplasm,
atypical, suspicious for malignancy and malignant resulted in a stratified risk of malignancy (ROM) (respective ROM – 6%, 53%, 79% and 100%). More recently, it was shown that designation of salivary aspirates as suspicious for malignancy had a similarly high ROM (84.2%) with no statistically significant difference between five international tertiary medical centers. In contrast, another abstract by the same authors reported a wide range in the ROM for salivary gland aspirates designated atypical (mean ROM 62.99%, range 0–73.08%).

The Milan System is a recently proposed scheme from an international group of cytologists with the support of the American Society of Cytopathology and International Association of Cytopathology that hopes to standardize the reporting of salivary gland FNA. The proposed system uses fairly traditional cytologic categories such as non-diagnostic, non-neoplastic, atypical, positive for neoplasm (benign or uncertain malignant potential), suspicious for malignancy, and malignant. One addition in the proposed Milan System divides neoplastic aspirates into benign neoplasm and neoplasm of uncertain malignant potential. This nuance shows acceptance of the importance of determining the presence of neoplasia and the significant overlap between benign and malignant neoplasms that exist in the setting of salivary glands. A recent comprehensive review examining the use of the proposed Milan categories in the literature found a risk stratification that may be useful to guide clinical management in patients with salivary gland tumors.

Another recent proposal uses a morphologic pattern-based approach, once the findings are sufficient for a neoplastic diagnosis, to stratify ROM, limit differential diagnostic considerations, and guide ancillary testing. Briefly, this pattern-based schema separates basaloid tumors from oncocytoid tumors (Table 1, first column). Basaloid tumors are further divided based on the quality of the stroma, such that those neoplasms lacking clearly identifiable and uniformly fibrillar stroma have a higher ROM than those with definitive and abundant fibrillar stroma occupying more than three-fourths of the material. On the oncocytoid side, the presence of mucus in the background or tumor cells with granular/vacuolated cytoplasm is associated with a higher ROM. Of note, this pattern-based risk stratification scheme could easily be used to supplement the proposed Milan System. This could be especially useful for salivary tumors categorized as “neoplasm of uncertain malignant potential” for which a broad ROM might be expected. Further separation of such aspirates using pattern-based features of stromal characteristics and background material could narrow the ROM and limit differential diagnostic considerations.

The purpose of the current study is to examine the interobserver variability of the proposed pattern-based risk stratification scheme of Griffith et al. Additionally, the ROM is also examined for each of the diagnostic categories to determine if the results corroborate those of the original proposal. Based on these data, the potential to reduce the number of pattern-based categories and simplify this system was also investigated.
Methods

Case selection

A total of 50 salivary gland aspirates were selected from cases between 2013 and 2015, divided evenly between two large academic centers, and examined by four cytopathologists from two institutions (Emory University and University of Pittsburgh Medical Center). Slides from salivary gland aspirates during the two year period were reviewed (CCG for Emory and SEM for UPMC) to select 25 cases from each institution. Cases with slides having material representative of the final prospective cytologic diagnosis were selected from cases and these slides were used for digital scanning. The cases included a variety of diagnoses from non-neoplastic to benign and malignant neoplasms, of which the malignant cases had confirmatory surgical follow-up. Histologic diagnoses of these cases with available surgical follow-up are detailed in Table 2. When available one representative slide each of Papanicolaou and Romanowsky stained smears and one hematoxylin and eosin (H&E) stained cell block section were chosen for digital scanning with an Aperio scanner (Leica Biosystems). 33 cases had one slide each of a Papanicolaou stained smear, Romanowsky stained smear and H&E stained cell block section, 13 had a Papanicolaou stained smear and Romanowsky stained smear, 3 had a Papanicolaou stained smear and an H&E stained cell block section and a single case had only a Papanicolaou stained smear. There was a two week washout period whereby slides were not reviewed until two weeks after selection. All cases were randomly coded with an alphanumeric code and digitally scanned slides were then evaluated independently by all authors using Aperio’s Imagescope viewer software (Leica Biosystems). Each case was evaluated for adequacy and categorized as non-neoplastic, atypical, neoplastic, or malignant. In cases considered at least suspicious for neoplasm, a pattern was assigned using the pattern-based risk stratification scheme proposed by Griffith et al (Table 1, first column). One author (CCG) was involved in the original development and proposal of this pattern based scheme and the other observers reviewed the original proposal article and a summary chart describing the pattern based classification for those lesions deemed at least neoplastic; however, no specific training was given in the recognition of these patterns. In order to examine the effect of simplification in this proposed pattern-based system, several categories were condensed on secondary analysis (Table 1, second column). In this simplified scheme, basaloid neoplasms were divided into those with fibrillary stroma (includes an aggregate of pleomorphic adenoma and basaloid neoplasms with fibrillary stroma) and those with non-fibrillary stroma (includes an aggregate of those neoplasms with hyaline stroma and mixed stroma). Similarly, oncocytoid neoplasms with cystic or other background were condensed into a single group.

Calculation of aggregate risk of malignancy

Aggregated ROM was calculated for each pattern by scoring each individual classification amongst the four observers as separate observations and using final histologic diagnoses from surgically excised cases when available. Chronic sialadenitis was defined as non-neoplastic. Benign neoplasms included PA, WT, and basal cell adenoma (BCA). Malignant tumors included MEC, epithelial-myoepithelial carcinoma (EMCA), basal cell adenocarcinoma (BCAC), sebaceous carcinoma, salivary duct carcinoma (SDC), AdCC and metastatic squamous cell carcinoma.
Determination of interobserver agreement

Due to the small number of cases and number of potential classifications precluding statistical analysis, interobserver agreement was determined using a scaled scoring scheme as outlined in Table 3. This study was meant to determine the interobserver variability of the proposed pattern-based scheme and therefore only cases with at least two observers classifying an aspirate as neoplastic were included for analysis. A total of four cases included in this analysis were not considered diagnostic of neoplasm by all authors (2 cases by one observer and 2 cases by two observers).

Results

Aggregated risks of malignancy

Table 4 shows the aggregated ROM for the proposed categories. Overall, there was a stratification in the ROM seen within basaloid neoplasms. Those tumors classified as “pleomorphic adenoma” and “basaloid neoplasm with fibrillary stroma” (Figure 1A) had a low ROM of 14.3% (3 of 21 and 2 of 14, respectively) on surgical follow-up. In comparison, basaloid neoplasms with non-fibrillary stroma (Figure 1B) had a higher ROM – 83.3% (5 of 6) and 60% (6 of 10) for those classified as having hyaline or mixed/other stroma, respectively. Cases classified as “pleomorphic basaloid neoplasm” also had a high ROM at 75% (3 of 4).

Condensation of these patterns into “basaloid neoplasm with fibrillary stroma” (to include “pleomorphic adenoma” and “basaloid neoplasm with fibrillary stroma”), “basaloid neoplasm with non-fibrillary stroma” (to include “basaloid neoplasm with hyaline stroma” and “basaloid neoplasm with mixed/other stroma”), and “pleomorphic basaloid neoplasm” allows simplification of the proposed scheme and maintains risk stratification with ROM being 14.3% (5 of 35), 68.8% (11 of 16), and 75% (3 of 4) for these respective groups.

Oncocytoid neoplasms also show stratification in ROM (Table 4). Cases classified as “Warthin tumor” had the lowest ROM of 0%, with all 16 such cases being benign on follow-up. ROM is progressively increased in the other categories when divided by background and cytoplasmic features from 25% (1 of 4) for “oncocytoid neoplasm with cystic background,” 40% (2 of 5) for “oncocytoid neoplasm with other background,” 80% (24 of 30) for “oncocytoid neoplasm with mucinous background,” (Figure 2) and 81.3% (13 of 16) for “oncocytoid neoplasm with granular/vacuolated cytoplasm.” Cases categorized as “pleomorphic oncocytoid neoplasm” were universally malignant (100% ROM) on follow-up (17 of 17).

As with the basaloid group, some condensation of these oncocytoid patterns is also possible. “Warthin tumor” can be maintained as a separate pattern, as the ROM in this study and the prior proposal was 0% (0 of 45) indicating that the diagnosis of WT can be made confidently when suitable diagnostic features are present. Such a category could allow some patients to more easily avoid surgery, depending on the clinical situation. Oncocytoid neoplasms with cystic or other background can be combined into a low risk oncocytoid neoplasm category and termed “oncocytoid neoplasm with cystic/other background.” The resultant ROM for this group is 33.3% (3 of 9), falling between the lowest risk group of
“Warthin tumor” and higher risk groups of “oncocytoid neoplasm with mucinous background” and “oncocytoid neoplasm with granular/vacuolated cytoplasm.”

**Interobserver agreement**

Of the 50 cases included in the study, 42 (84%) were considered neoplastic by at least two observers and were therefore included for the subsequent analysis of interobserver variability. Overall, interobserver agreement was quite good, despite the high number of potential categories in the initial proposal of Griffith et al (Figure 3A). Using the scoring detailed in Table 3, 23.8% (10 of 42) had uniform agreement, 33.3% (14 of 42) majority agreement, and 11.9% (5 of 42) divided agreement. More than half of the cases (57.1%; 24 of 42) had uniform or majority agreement such that the ROM and differential diagnostic considerations would be similar despite minor differences in final classification. Only 21.4% (9 of 42) and 9.5% (4 of 42) had minimal or no agreement, respectively.

When the more simplified aforementioned classification scheme was applied, there was an increase in the number of cases with uniform agreement and a decrease in cases with no agreement (Figure 3B). With this simplified scheme, half (n=7) of the cases previously scored as having majority agreement resulted in uniform agreement such that 40.5% (17 of 42) aspirates had uniform agreement. All seven of these cases with a shift in agreement to uniform were due to the condensation of basaloid neoplasms into those with fibrillary or non-fibrillary stroma. The number of cases with uniform or majority agreement, however, remained the same (57.1%).

With condensation of diagnostic categories, only a single case improved from no agreement to minimal agreement. This was a case of basal cell adenoma that demonstrated improved agreement due to the condensation of the basaloid categories. Additionally, a single case was raised from minimal agreement to divided agreement. Again, this improvement was due to condensation of basaloid categories and was seen in a case of pleomorphic adenoma.

Of note, three cases diagnosed as pleomorphic adenoma on surgical follow-up in our study that had low interobserver agreement as a result of categorization of fibrillary stroma as background mucus in an oncocytoid neoplasm (i.e. one observer classified an aspirate from a PA as an oncocytoid neoplasm with mucinous background while the others agreed on the designation of PA) (Figure 4).

**Discussion**

In this study, we examine the interobserver variability of the proposed pattern-based risk stratification scheme of Griffith et al using digital slides from two institutions. We also confirm the ability of this pattern-based approach to stratify aspirates into groups with different ROM, which is similar to that of the initial proposal.

**Risk stratification with this pattern-based scheme is confirmed**

Due to the small number of cases included in this study, aggregated ROM was calculated for each categorization made by individual observers. Using these aggregated ROM, the current study shows a similar risk stratification profile to that reported in the original proposal, and

*Cancer*; Author manuscript; available in PMC 2018 October 01.
some categories show nearly identical risks of malignancy.\textsuperscript{12} Specifically, pattern categories with nearly identical ROM include the basaloid neoplasm with fibrillary stroma (14.3% compared to 15.4%), basaloid neoplasm with mixed/hyaline stroma (60% in both studies), Warthin tumor (0% in both studies), oncocytoid neoplasm with mucinous background (80% in both studies), oncocytoid neoplasm with granular/vacuolated cytoplasm (81.3% compared to 84.6%) and pleomorphic oncocytoid neoplasm (100% in both studies).

Several pattern categories had considerably different risks of malignancy in this study compared to the prior publication, including the basaloid neoplasm with hyaline stroma (83.3% compared to 42.9%) and oncocytoid neoplasm with other background (40% compared to 4%). Some of this difference may be attributed to the small number of cases included in these pattern categories resulting in greater fluctuations in ROM between studies. This fluctuation in the ROM for these categories also prompted the condensation of these categories to simplify the scheme which could resolve some of these differences.

There are several limitations in the ROM calculates presented in this study which deserve discussion. The ROM which are calculated in this study as in the initial proposal are biased due to selection of a subset of salivary gland tumor cytology cases and the actual ROM for these different pattern-based categories would likely be lower in a prospective study or in a series looking at all salivary gland aspirates over a given time period. The lack of surgical follow-up in some cases could also introduce bias into the ROM calculations. Finally, the small number of cases included in this study may not reflect the risk in a larger series although it should be noted that the current series shows similar ROM as previously discussed. Despite these limitations, the findings do suggest that the different patterns have the potential to result in risk stratification for salivary gland tumors, as shown in the prior initial larger study based on histological follow-up.\textsuperscript{12} Additional studies will be needed to more accurately define the ROM in actual practice.

This study shows that the proposed pattern-based risk stratification scheme proposed by Griffith et al for the classification of salivary gland aspirates has good interobserver agreement. These results were attained despite the fact that this pattern-based system incorporates more categories than most traditional risk categorization schemes in cytopathology such as the Bethesda systems for reporting cervicovaginal and thyroid cytologic samples. The reasons for this overall good interobserver variability are likely related to the fact that this proposed system is based on simple morphological distinguishing features, many of which are already considered by practicing cytopathologists when presented with a salivary gland tumor aspirate.

The ability to identify a salivary gland aspirate as a basaloid neoplasm offers a general differential diagnosis that many experienced cytopathologists recognize. By incorporating stromal characteristics of basaloid neoplasms in the proposed pattern-based approach, differential diagnostic considerations are refined and ROM for these aspirates is stratified, as was shown in the original scheme proposal.\textsuperscript{9}

In contrast to the basaloid group, the concept of an oncocytoid neoplasm in aspiration cytology of the salivary glands is not as well established. Despite this, the ability to add risk

\textit{Cancer}. Author manuscript; available in PMC 2018 October 01.
stratification and limit differential diagnostic considerations based on background, cytoplasm, and/or cytonuclear grade, in the setting of such oncocytoid neoplasms is confirmed by this study. As with basaloid neoplasms, the features used to risk stratify oncocytoid neoplasms include well-known cytomorphologic features. For example, the presence of granular cytoplasm is a well-established feature of AciCC. Although secretory carcinoma classically displays more vacuolated cytoplasm, similarities between AciCC and secretory carcinoma make the distinction of these two entities difficult on aspiration cytology. Categorization of aspirates as oncocytoid neoplasm with granular/vacuolated cytoplasm is designed to aid in the recognition of these two cytologically bland malignancies and highlights the significant ROM associated with this pattern.

Perhaps one of the more difficult differential diagnoses in salivary gland cytopathology is the distinction of WT with mucinous metaplasia from a low- to intermediate-grade MEC. Aspirates from both of these entities show low-grade oncocytoid epithelial cells with or without mucus cells and frequently have a prominent mucinous background. Although a prior study showed that the presence of mucus was not a significant distinguishing feature between WT and MEC,\textsuperscript{14} our findings suggest that the presence of prominent mucus indicates a higher ROM on histologic follow-up. Whereas some cases classified as oncocytoid neoplasm with mucinous background were benign on excision (mostly WT), these findings demonstrate the importance of this categorization to highlight the significant ROM associated with this pattern. In the absence of a risk stratification scheme, such an aspirate would often be diagnosed descriptively with a differential diagnosis including WT and MEC; however, the significant ROM may not be communicated effectively to the treating physician. In such cases, categorization as positive for an oncocytoid neoplasm with mucinous background could be important to prompt surgical excision.

Simplification of the scheme with condensation of similar categories preserves risk stratification and offers some improvements in interobserver variability

Although there was good agreement with the initially proposed pattern-based system, we felt that there was room for simplification through condensation of categories with overlapping morphologic features, differential diagnostic considerations, and ROM. Hence, the number of categories was reduced and, not unexpectedly, the overall interobserver variability was improved. Basaloid neoplasms with hyaline and mixed/other stroma were conveniently condensed into a single category of basaloid neoplasm with non-fibrillar stroma. This new grouping has a high ROM and the distinction between subtly different forms of stroma was one source of interobserver disagreement. As seen in Table 4, almost all of the improvements in agreement were due to the condensation of basaloid neoplasm categories. Less condensation of oncocytoid groups was possible due to the fact that this group of tumors exhibit more divergent morphologic features. Additionally, the condensation of oncocytoid neoplasm with cystic background or other background into a single category had minimal effect in improving interobserver agreement. Further improvements in interobserver variability may be possible with a training set prior to evaluating salivary gland aspirates.
Traditional pitfalls in salivary gland aspiration cytology must still be avoided

There are several traditional pitfalls in salivary gland cytology that must still be remembered for those attempting to employ this pattern-based risk stratification scheme, or another proposed scheme. For example, one of the difficulties in the diagnosis of AcIC is the potential to misinterpret neoplastic acinar cells as normal non-neoplastic acinar tissue. Features that are useful in recognizing AcIC are the finding of exclusively acinar cells without normal architectural clustering and a lack of ductal cells, cells with reduced cohesion, and cells with fragile cytoplasm indicated by naked nuclei.

Difficulty in differentiating the fibrillary stroma of PA from mucus can also cause diagnostic confusion between aspirates from PA and MEC. In this setting, histiocytes within mucus can appear to be myoepithelial cells in myxoid stroma and vice versa. Three cases in our study that had low interobserver agreement were as a result of this pitfall. In each of these three cases a single observer classified an aspirate from a PA as an oncocytoid neoplasm with mucinous background while the others agreed on the designation of PA. This potential pitfall may have been accentuated by the use of only limited numbers of representative slides and the use of digitally scanned slides with limited resolution.

Limitations

Nearly a third of cases in this series that had no to minimal agreement (12–13 of 42 cases [28.6–31.0%]) which raise questions about the clinical utility of the proposed classification scheme. However, there is currently no system for the classification of salivary gland tumor aspiration cytology and this it is generally accepted that this is a challenging area because of the morphologic overlap seen in these tumors spanning both benign and malignant behavior. Even in well-established and accepted standard cytology categorization systems such as The Bethesda System for Reporting Cervical Cytology there is significant variability, especially in the use of ASCUS with one study showing that of 1473 monolayer cytology specimens originally designated ASCUS, only 633 (43.0%) were interpreted as ASCUS on second review. Other categories showed agreement on second review of 77.6% for negative, 68.0% for LSIL and 47.1% for ≥HSIL. Interobserver variability has also been shown in thyroid cytopathology in a variety of studies. While these prior studies are not directly comparable to the results in our study they highlight the inherent interobserver variability that can be expected in categorization schemes for cytologic samples. Given the inherent interobserver variability of classification schemes, the complexity of salivary gland tumor cytology, and the early phase of this proposed scheme, this poor agreement on these cases should not preclude further work to develop a standardized system for reporting salivary gland cytology. Continued refinement of this pattern-based system and increased experience might further improve interobserver agreement in improve the clinical potential of such a system.

Some of the limitation in the current study may also have contributed to these cases with poor interobserver variability and include the lack of specific training and/or experience with this specific pattern-based risk stratification scheme. As users become more familiar with this system, there may be a potential to see improved agreement amongst observers. We also have not examined the source of the lack of agreement in these cases specifically but given
the fairly even distribution of histologic diagnoses across the areas of agreement as seen in Figure 3, it may be more related to the quality or cellularity of the available slides than the specific diagnosis or pattern.

**Conclusion**

This study adds further support to the use of a pattern-based risk stratification scheme in the classification of salivary gland tumors. The ability to stratify the ROM using the proposed patterns confirms similar results of the initial Griffith proposal. Additionally, despite the complexity of the proposed scheme we show that interobserver variability is good and can be improved with condensation of some of these categories, while still preserving the desired risk stratification and differential diagnoses. Further study is planned to examine a more realistic prediction of malignancy risk for these various patterns in a more systematic examination of salivary gland aspirates. This pattern and risk-based classification scheme has a potential to be used in conjunction with the forthcoming Milan System for Reporting Salivary Gland Cytopathology. The proposed Milan System could serve as the primary diagnostic categorization (e.g. positive for neoplasm of uncertain malignant potential) with this pattern-based scheme serving as a secondary or free text diagnosis (e.g. basaloid neoplasm with non-fibrillary stroma). By clarifying the pattern seen in the aspirate, a more granular impression of the ROM can be given and therefore be more helpful in determining the management for patients with salivary gland tumors, akin to The Bethesda System for Reporting Thyroid Cytology.

**Acknowledgments**

Research reported in this publication was supported in part by the Cancer Tissue and Pathology Shared Resource of Winship Cancer Institute of Emory University and NIH/NCI under award number P30CA138292. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Funding sources: None.

**References**


Figure 1. Basaloid neoplasms
A: Aspirate showing basaloid cells with very scant stroma which was classified by two observers as fibrillary (Diff Quik, 400×). One observer classified the case as a basaloid neoplasm with mixed/other stroma. Surgical follow-up showed pleomorphic adenoma. B: Aspirate with non-fibrillary stroma (Diff Quik, 400×) that was classified as a basaloid neoplasm with hyaline stroma by three observers and as mixed other stroma by one observer. Surgical excision showed adenoid cystic carcinoma.
Figure 2. Oncocytoid neoplasm with mucinous background

A: This aspirate was classified as an oncocytoid neoplasm with mucinous background by all four observers (Diff Quik, 200×). The excised tumor was diagnosed as a low grade mucoepidermoid carcinoma. B: Only two observers classified this aspirate as neoplastic and both classified it as an oncocytoid neoplasm with mucinous background (Diff Quik, 200×). Surgical excision showed a Warthin tumor with mucinous metaplasia.
Figure 3.
Interobserver agreement using (A) the original pattern-based classification scheme proposed by Griffith et al or (B) the newer simplified pattern-based classification scheme for salivary gland FNA. Abbreviations: AdCC – adenoid cystic carcinoma, BCA – basal cell adenoma, BCACA – basal cell adenocarcinoma, EMCA – epithelial myoepithelial carcinoma, Infl – inflammatory lesion, MEC – mucoepidermoid carcinoma, PA – pleomorphic adenoma, SC – secretory carcinoma, SDC – salivary duct carcinoma, SDCexPA – salivary duct carcinoma ex pleomorphic adenoma, SqCC – squamous cell carcinoma, Unk – unknown diagnosis, WT – Warthin tumor. (+1) denotes cases in which there is no histologic follow-up but the clinical
history, original sign out and consensus diagnoses from the current review support designation as PA or an inflammatory lesion.
Figure 4.
Pitfall demonstrating misinterpretation of fibrillary stroma of a pleomorphic adenoma as a mucinous background in an oncocytoid neoplasm. This aspirate from a pleomorphic adenoma shows fibrillary stroma with embedded myoepithelial cells (Diff Quik, 200×). However, it is possible to interpret the myoepithelial cells as oncocytoid epithelial cells floating in smeared mucus. This pitfall can result in a higher risk of malignancy being assigned to a benign tumor if not recognized.
Table 1
Original Pattern Based Salivary Aspirate Risk Stratification Categories and Proposed Simplification

<table>
<thead>
<tr>
<th>Original Proposal</th>
<th>Simplified Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Basaloid Neoplasms</strong></td>
<td></td>
</tr>
<tr>
<td>Pleomorphic adenoma</td>
<td>Basaloid Neoplasms</td>
</tr>
<tr>
<td>Basaloid neoplasm with fibrillary stroma</td>
<td>Basaloid neoplasm with fibrillary stroma</td>
</tr>
<tr>
<td>Basaloid neoplasm with hyaline stroma</td>
<td>Basaloid neoplasm with non-fibrillary stroma</td>
</tr>
<tr>
<td>Basaloid neoplasm with mixed/other stroma</td>
<td></td>
</tr>
<tr>
<td>Pleomorphic basaloid neoplasm</td>
<td>Pleomorphic basaloid neoplasm</td>
</tr>
<tr>
<td><strong>Oncocytoid Neoplasms</strong></td>
<td></td>
</tr>
<tr>
<td>Warthin tumor</td>
<td>Warthin tumor</td>
</tr>
<tr>
<td>Oncocytoid neoplasm with cystic background</td>
<td>Oncocytoid neoplasm with cystic/other</td>
</tr>
<tr>
<td>Oncocytoid neoplasm with other background</td>
<td></td>
</tr>
<tr>
<td>Oncocytoid neoplasm with mucinous background</td>
<td>Oncocytoid neoplasm with mucinous background</td>
</tr>
<tr>
<td>Oncocytoid neoplasm with granular/vacuolated cytoplasm</td>
<td>Oncocytoid neoplasm with granular/vacuolated cytoplasm</td>
</tr>
<tr>
<td>Pleomorphic oncocytoid neoplasm</td>
<td>Pleomorphic oncocytoid neoplasm</td>
</tr>
<tr>
<td>Category</td>
<td>Subcategory</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>--------------------------------------</td>
</tr>
<tr>
<td>No Histologic follow-up (11)</td>
<td></td>
</tr>
<tr>
<td>Non-neoplastic (1)</td>
<td>Chronic sialadenitis (1)</td>
</tr>
<tr>
<td>Benign neoplasm (17)</td>
<td>Pleomorphic adenoma (9)</td>
</tr>
<tr>
<td></td>
<td>Warthin tumor (7)</td>
</tr>
<tr>
<td></td>
<td>Basal cell adenoma (1)</td>
</tr>
<tr>
<td>Low grade malignant neoplasm (10)</td>
<td>Mucoepidermoid carcinoma (4 LG, 2 IG)</td>
</tr>
<tr>
<td></td>
<td>Epithelial-myoepithelial carcinoma (2)</td>
</tr>
<tr>
<td></td>
<td>Basal cell adenocarcinoma (1)</td>
</tr>
<tr>
<td></td>
<td>Secretory carcinoma (1)</td>
</tr>
<tr>
<td>Hi Grade malignant neoplasm (11)</td>
<td>Salivary duct carcinoma (5)</td>
</tr>
<tr>
<td></td>
<td>Mucoepidermoid carcinoma (2 HG)</td>
</tr>
<tr>
<td></td>
<td>Adenoid cystic carcinoma (2)</td>
</tr>
<tr>
<td></td>
<td>Metastatic squamous cell carcinoma (2)</td>
</tr>
</tbody>
</table>

Abbreviations: LG - low grade; IG - intermediate grade; HG - high grade
Table 3

Scoring of agreement for cases for secondary interpretation

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Uniform agreement – all observers gave the same exact categorization OR if some authors (up to 2) considered the lesion non-neoplastic, the other authors gave the same exact category (i.e. if only one author considers the case neoplastic, the case is excluded).</td>
</tr>
<tr>
<td>2</td>
<td>Majority agreement – not all observers gave the same exact categorization but only two different categories were provided across all observers and these categories are very close in morphology and overall risk of malignancy (i.e. pleomorphic adenoma vs cellular basaloid neoplasm with fibrillary stroma).</td>
</tr>
<tr>
<td>3</td>
<td>Divided agreement – up to three different categorizations are provided and all are at least basaloid/oncocytoid and there is a close similarity in risk of malignancy for those provided categories (i.e. Warthin tumor, oncocytoid neoplasm with cystic background and oncocytoid neoplasm with other background). Also, cases with three observers in complete agreement but with a single outlier observer providing a categorization with drastically different risk of malignancy (i.e. basaloid neoplasm with fibrillary stroma x3; oncocytoid neoplasm with mucinous background x1).</td>
</tr>
<tr>
<td>4</td>
<td>Minimal agreement – two or three different categorizations are provided and these may span both basaloid and oncocytoid and/or a wider range of malignancy (i.e. Warthin x2; Oncocytoid neoplasm with cystic background x1; Oncocytoid neoplasm with mucinous background x1).</td>
</tr>
<tr>
<td>5</td>
<td>No agreement – all observers provided a different categorization and these range from both basaloid and oncocytoid and/or very low risk of malignancy to very high risks of malignancy.</td>
</tr>
</tbody>
</table>
Table 4
Aggregated Risks of Malignancy by Proposed Pattern Based Categorizations

<table>
<thead>
<tr>
<th>Pattern Category</th>
<th>Original Proposal</th>
<th>Aggregated ROM, % (n)</th>
<th>Simplified Classification</th>
<th>Aggregated ROM, % (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pleomorphic adenoma</td>
<td>14.3% (3/21)</td>
<td>Basaloid neoplasm</td>
<td>14.3% (5/35)</td>
<td></td>
</tr>
<tr>
<td>Basaloid neoplasm with fibrillary stroma</td>
<td>14.3% (2/14)</td>
<td>Basaloid neoplasm</td>
<td>14.3% (5/35)</td>
<td></td>
</tr>
<tr>
<td>Basaloid neoplasm with hyaline stroma</td>
<td>83.3% (5/6)</td>
<td>Basaloid neoplasm</td>
<td>68.8% (11/16)</td>
<td></td>
</tr>
<tr>
<td>Basaloid neoplasm with mixed/other stroma</td>
<td>60% (6/10)</td>
<td>Basaloid neoplasm</td>
<td>68.8% (11/16)</td>
<td></td>
</tr>
<tr>
<td>Pleomorphic basaloid neoplasm</td>
<td>75% (3/4)</td>
<td>Pleomorphic basaloid</td>
<td>75% (3/4)</td>
<td></td>
</tr>
<tr>
<td>Warthin tumor</td>
<td>0% (0/16)</td>
<td>Warthin tumor</td>
<td>0% (0/16)</td>
<td></td>
</tr>
<tr>
<td>Oncocytoid neoplasm with cystic background</td>
<td>25% (1/4)</td>
<td>Oncocytoid neoplasm</td>
<td>33.3% (3/9)</td>
<td></td>
</tr>
<tr>
<td>Oncocytoid neoplasm with other background</td>
<td>40% (2/5)</td>
<td>Oncocytoid neoplasm</td>
<td>33.3% (3/9)</td>
<td></td>
</tr>
<tr>
<td>Oncocytoid neoplasm with mucinous background</td>
<td>80% (24/30)</td>
<td>Oncocytoid neoplasm</td>
<td>80% (24/30)</td>
<td></td>
</tr>
<tr>
<td>Oncocytoid neoplasm with granular/vacuolated cytoplasm</td>
<td>81.3% (13/16)</td>
<td>Oncocytoid neoplasm</td>
<td>81.3% (13/16)</td>
<td></td>
</tr>
<tr>
<td>Pleomorphic oncocytoid neoplasm</td>
<td>100% (17/17)</td>
<td>Pleomorphic oncocytoid neoplasm</td>
<td>100% (17/17)</td>
<td></td>
</tr>
</tbody>
</table>