Layfield Lester J. (Orcid ID: 0000-0003-1344-6506)
Esebua Magda (Orcid ID: 0000-0001-5884-9603)
Pantanowitz Liron (Orcid ID: 0000-0001-8182-5503)
Maleki Zahra (Orcid ID: 0000-0003-3273-4802)
Cantley Richard (Orcid ID: 0000-0002-6564-3889)

Salivary Gland Neoplasms with Basaloid Features in the Era of the Milan System for Reporting Salivary Gland Cytology: Classification and Interobserver Agreement

Lester J. Layfield, MD[*], Magda Esebua, MD[*], Liron Pantanowitz, MD, MHA[+], Zahra Maleki, MD, FCAP, MIAC[#], Maryna Vazmitsel, MD[**], Zubair Baloch, MD, PhD[++]., Richard L. Cantley, MD[##], Robert Schmidt, MD, PhD[°]

[*]Department of Pathology and Anatomical Sciences
University of Missouri, Columbia, Missouri
[+]Department of Pathology
University of Michigan, Ann Arbor, Michigan
[#]Department of Pathology
Johns Hopkins Hospital, Baltimore, Maryland
[**]Department of Medical Pathology and Laboratory Medicine
UC-Davis Health, Sacramento, California
[++]Department of Pathology and Laboratory Medicine
Hospital of the University of Pennsylvania
[##]Department of Pathology
University of Michigan, Ann Arbor, Michigan
[°]Department of Pathology
University of Utah Health Care and ARUP Laboratories, Salt Lake City, Utah

Lester J. Layfield -layfieldl@health.missouri.edu
Magda Esebua – esebuam@health.missouri.edu
Liron Pantanowitz – pantanowitzl@upmc.edu
Zahra Maleki – zmaleki1@jhmi.edu
Maryna Vazmitsel – mvazmitsel@ucdavis.edu
Zubair Baloch – baloch@mail.med.upenn.edu
Richard L. Cantley – rcantley@med.umich.edu
Robert Schmidt – Robert.Schmidt@hsc.utah.edu

***There are no conflicts of interest**
****There are no grants or any other funding**

Key Words: Milan System, Pleomorphic adenoma, Monomorphic adenoma, Adenoid Cystic carcinoma

Running Title: Interobserver Agreement for Basaloid neoplasms of Salivary Gland

Abstract:

Background: The Milan System for Reporting Salivary Gland Cytopathology (MSRSGC) has been shown to have moderate to good reproducibility for categorization of salivary gland fine-needle aspiration (FNA) specimens. Less is known of its accuracy and interobserver reproducibility for categorization of the diagnostically difficult group of basaloid neoplasms.

Methods: Forty-five salivary gland specimens with a basaloid morphology (pleomorphic and monomorphic adenomas and adenoid cystic carcinomas) were independently assigned by seven cytopathologists to one of the MSRSGC categories. Interobserver agreement was assessed for average agreement, chance expected agreement and by Cohen's kappa and diagnostic accuracy. Correlation of the salivary gland neoplasm of unknown malignant potential (SUMP) category with histologic diagnosis and benign or malignant designation along with interobserver reproducibility were calculated.

Results: Average observed agreement for assignment to the MSRSGC was 46% and Cohen's kappa = 0.2%. The SUMP category did not correlate with tumor type or with the benign or malignant nature of the neoplasm. Diagnostic specificity and sensitivity were 92 and 100% for consensus diagnosis, but were 76 and 77% for individual diagnoses..

Conclusion: The interobserver agreement in categorizing basaloid neoplasms by the MSRSGC is poorer than for salivary gland lesions overall. This reflects the difficulty in diagnosing basaloid neoplasms. Nonetheless, diagnostic accuracy appears similar to that of salivary gland neoplasms as a whole.

Introduction:

Fine-needle aspiration cytology (FNA) has been highly successful in the diagnosis of both neoplastic and non-neoplastic lesions of the salivary glands[1-3] Despite the overall high accuracy and utility of the technique, some diagnostic problems have been recognized.[4-15] Prominent among the diagnostic problems is the group of neoplasms characterized by a small "basaloid" cell morphology and variable amounts of stroma. The commonly evaluated salivary gland neoplasms in this group include of cellular pleomorphic adenomas, monomorphic adenomas, basal cell adenomas and adenoid cystic carcinomas.[5,7-10] A number of authors have reviewed the basal cell adenomas, criteria for the cytologic diagnosis of adenoid cystic carcinomas[14,15] and other basaloid neoplasms[16]. Despite the published diagnostic criteria, distinction of cellular pleomorphic adenomas, basal cell adenomas and monomorphic adenomas from some cases of adenoid cystic carcinomas and other basaloid neoplasms remains diagnostically challenging. This difficulty in distinction of benign neoplasms with basaloid features from adenoid cystic carcinomas decreases the overall utility of FNA for the separation of benign and malignant neoplasms of the salivary glands.

The Milan System for Reporting Salivary Gland Cytopathology (MSRSGC) was developed to offer a reproducible classification system for cytologic specimens of the salivary glands obtained by fine-needle aspiration cytology (FNA) along with recommendations for subsequent patient follow-up.[17] The system acknowledges the diagnostic difficulty in definitive separation of a subset of benign from malignant neoplasms of the salivary glands and has developed a set of categories to address this issue. These categories include: neoplasm, suspicious for malignancy, and malignant. The neoplastic category

is subdivided into benign neoplasm and salivary gland neoplasm of uncertain malignant potential (SUMP). The utility of the SUMP category to which cytopathologists can assign neoplasms for which they are unsure as to whether the case they are evaluating is definitely benign, suspicious for malignancy or malignant diagnosis is unclear. Follow-up recommendations for the SUMP category are distinct from those given for the benign neoplasm, suspicious for malignancy and malignant categories.

The SUMP category was designed to improve the diagnostic accuracy of the neoplasm, benign category and the suspicious for malignancy and malignant categories by placing neoplasms especially difficult to categorize as benign or malignant into the indeterminate SUMP category. By removing specimens which are especially difficult to classify as benign or malignant from the more specific categories, it was hoped that the benign and the combined suspicious for malignancy and malignant categories would more accurately classify benign and malignant neoplasm with few false positive and false negative classifications.

We investigated if the use of a SUMP category resulted in good sensitivity and specificity for the recognition of salivary gland malignancies in the diagnostically difficult category of neoplasms characterized by a small basaloid morphology and variable amounts of stroma. We studied whether or not the categories benign and malignant retained their high diagnostic accuracy and predictive value even when exclusively small basaloid neoplasms were studied. We investigated if the use of the SUMP category was associated with good accuracy of assignment of basaloid neoplasms to the neoplasm, benign or malignant categories. We also investigated how accuracy of category assignment for the diagnostically challenging basaloid neoplasms compares to the accuracy of assignment of unselected salivary gland neoplasms as reported in the literature by comparing malignancy risks for the diagnostic categories. In addition, we evaluated the reproducibility of the Milan categories among cytopathologists for the assessment of monomorphic adenomas, cellular pleomorphic adenomas and adenoid cystic carcinomas.

Methods and Materials

Following approvals by the Institutional Review Boards at the University of Missouri and the University of Michigan, an electronic search of the cytopathology records at each institution was undertaken for all cases of pleomorphic adenoma, monomorphic adenoma, basal cell adenocarcinoma, adenoid cystic carcinoma and polymorphous low-grade adenocarcinoma. A total of forty-five cases met the search criteria and had adequate smears for evaluation. Each of these cases had the surgical pathology files searched for the corresponding excision specimens and the cytologic diagnoses were correlated with the subsequent histopathologic diagnoses. For the purposes of final diagnosis, the surgical pathology diagnosis was used to determine if a neoplasm was benign or malignant and establish the precise histologic type. The slides were independently reviewed by seven cytopathologists, all but one of whom were board certified cytopathologists. The single non-board-certified cytopathologist had approximately 10 years of experience as a surgical pathologist with interest in head and neck pathology and a similar length of experience with the cytopathology of head and neck lesions. Each cytopathologist had between 4 and 35-years experience in evaluating FNA specimens obtained from the head and neck. Each cytopathologist independently reviewed the slides without prior knowledge of either the cytologic diagnosis of record or the associated surgical pathology diagnosis. Each slide was assigned to one of the Milan System categories (non-diagnostic, non-neoplastic, atypia of undetermined significance (AUS), neoplasm benign, salivary gland neoplasm of uncertain malignant potential (SUMP), suspicious for malignancy or malignant). Cytologic diagnoses were correlated with the final surgical pathology diagnosis. For each case, a consensus diagnosis was obtained and defined as the majority diagnosis (agreement between at least 4 of 7 cytopathologists).

Malignancy risk was calculated for the consensus categories neoplasm benign, SUMP, suspicious for malignancy and malignant as well as for the combined category composed of the categories suspicious for malignancy and malignant. Malignancy risk was calculated as the number of malignancies in category/total cases in that category. A low malignancy risk in a benign category indicated a high diagnostic accuracy for recognition of benign neoplasms while a high malignancy risk correlated with a high diagnostic accuracy for a malignant categorization. Percentages of neoplasms assigned correctly to the benign or malignant categories were also calculated.

The predictive value of a negative test was calculated for the neoplasm benign category and the predictive value of a positive test was calculated for the malignant category.

Accuracy of assignment of basaloid neoplasms to the neoplasm benign and malignant categories was compared with accuracy of assignment of all salivary gland neoplasms reported in the literature.

The data was analyzed to determine if there was a statistical difference between the surgical pathology diagnoses for cases where a majority diagnosis existed vs. those where no majority diagnosis occurred. In the group of cases where a majority diagnosis was present, statistical analysis was preformed to see if there was a difference between histologic diagnoses for the groups, benign neoplasm, SUMP and malignant. Cases with a diagnosis of SUMP were correlated with subsequent histology to determine if there was a relationship between a SUMP diagnosis and the presence of a benign or malignant neoplasm and for type of neoplasm present.

Degree of agreement was calculated based on all categories (no partial credit given) with average observer agreement between pairs of observers being calculated along with expected agreement and Cohen's kappa statistic. Agreement based on two categories (benign vs. malignant) was calculated for average observed agreement between pairs of observers and for average kappa statistic. A chi square test was preformed to determine if diagnostic accuracy varied with experience.

Results

The 45 cases that underwent categorization had both malignant and benign diagnoses as determined by final surgical pathology review (Table 1). Forty-nine percent (22 of 45) were classified as malignant while 51% (23 of 45) were adenomas. Table 2 gives the reviewers' diagnoses and final surgical pathology diagnosis. A diagnosis of suspicious was given by at least 1 reviewer in only 30 of 315 categorizations and SUMP in 92 categorizations.

Ninety-one percent of benign basaloid neoplasms were assigned to the neoplasm benign category and 90% of malignant basaloid neoplasms were assigned to the malignant category. Ninety-one percent of basaloid malignancies were assigned to the combined category of suspicious for malignancy and malignant. The sensitivity and specificity of consensus assignment of basaloid neoplasms to the malignant category were 90% and 91% respectively when only the malignant category was considered a true positive and the sensitivity and specificity were 91% and 92% when consensus diagnoses of suspicious for malignancy and malignant categorizations were considered true positives.

Malignancy risks associated with assignment of basaloid neoplasms to the Milan System categories are shown in Table 3. The malignancy risk of the neoplasm benign category was 8% and was 90% for the malignant category and 91% for the combined suspicious for malignancy and malignant category. The malignancy risk for SUMP was 44%. The negative predictive value was 92% while the positive predictive value was 90%. Comparison of malignancy risks associated with categorization of basaloid neoplasms with those associated with published risks of malignancy for unselected salivary

gland neoplasms are shown in Table 3. The level of cytopathologists' experience did not correlate with diagnostic accuracy (Table 4).

To better correlate impact of case assignment to categories, consensus diagnoses were formulated as the diagnosis given by four or more reviewers for each specimen. In 13 (29%) cases, no consensus diagnosis was obtained but in the remaining specimens a consensus diagnostic category was obtained (Table 5). Distribution of consensus diagnostic categories had benign neoplasm as the most common consensus category (27%) while the majority of the remaining cases were nearly evenly distributed between the SUMP and malignant categories (20% and 22% respectively). Suspicious for malignancy was the consensus categorization in only a single case (Table 5). The case with a consensus categorization of suspicious for malignancy was an adenoid cystic carcinoma. Cases designated as SUMP were nearly equally divided between benign and malignant neoplasms. Table 6 tabulates the distribution of neoplasm types among cases with a consensus categorization of SUMP. There was no significant correlation between histologic tumor type and the category SUMP ($p=0.72$). Assignment to the SUMP category did not predict the benign or malignant nature of a specimen. No correlation of a SUMP diagnosis with a benign or malignant histologic diagnosis was found ($p=0.64$). The association between the consensus category SUMP and the benign or malignant surgical diagnosis of the sample reviewed is shown in Table 7. Table 8 shows the correlation of final histologic diagnosis with consensus categorization. No consensus was achieved for the case of basal cell carcinoma and the consensus category SUMP was given for the single case of polymorphous low-grade adenocarcinoma. Table 9 documents the interobserver agreement for the categories SUMP, benign and malignant. The interobserver agreement of the SUMP category was 88.8% with a kappa value of 0.28 (fair agreement). The interobserver agreements for the benign and malignant categories were significantly greater than for the SUMP category. Tables 10 and 11 document the diagnostic accuracy for separation of cases into

benign neoplasm and malignant using the consensus and the individual diagnoses. The sensitivity and specificity for the consensus diagnoses were both 92 and 100% respectively but 76% and 77% for the individual diagnoses.

Assignment to one of the Milan system categories demonstrated an average observed agreement between pairs of observers of 46% (range: 36-62%). The expected agreement was 25% and Cohen's kappa statistic was 0.2%. This corresponds to only fair agreement. When agreement was based on a two-category system (benign/malignant) average observed agreement between pairs of observers was 73% (range 62 to 93%). The expected agreement was 54% and the average kappa statistic was 0.41. This corresponds to a moderate level of agreement.

Case 42 demonstrates a high level of agreement for a benign diagnosis which was histologically confirmed as a cellular mixed tumor. The smears show a population of plasmacytoid to small basaloid cells lying in a background of a myxoid stroma (Figure 1). Case 28 had near uniform agreement for malignant diagnosis and was histologically shown to be an adenoid cystic carcinoma. The smears demonstrated fight clusters of small basaloid cells. Little stroma was present. Some cell groups had a "finger in glove" appearance (Figure 2). Case 40 lacked a majority diagnosis. Three reviewers assigned it to the SUMP category. Three reviewers designated it "suspicious for malignancy" and a single rater categorized it as a benign neoplasm. Histologically, it was a monomorphic adenoma, smears demonstrated cell clusters composed of relatively monomorphous cells. Single cells and small groups of polygonal cells exfoliated off the larger clusters (Figure 3).

Discussion:

A group of salivary gland neoplasms including monomorphic adenomas, cellular pleomorphic adenomas and adenoid cystic carcinomas has been recognized as difficult to distinguish cytologically.[7-13]

These neoplasms are characterized cytologically by tightly cohesive groups of relatively small cells with scant cytoplasm often associated with stromal material. This overlapping morphology can result in confusion between benign adenomas and some adenoid cystic carcinomas.[7-9]

The MSRSGC was designed to facilitate clinically useful classification of diagnostically difficult salivary gland FNA specimens.[17] Four categories defined by the MSRSGC are purportedly most useful in classifying salivary gland specimens characterized by tight clusters of relatively small cells associated with variable amounts of stroma. These categories are: 1) neoplasm benign, 2) neoplasm SUMP, 3) suspicious for malignancy, and 4) malignant. One of the values of standardized categorization systems is to improve interobserver consistency of diagnosis there by allowing for consistent and appropriate clinical management of patients. Our study investigated the diagnostic accuracy for separation of benign from malignant and the interobserver agreement for assignment of specimens with a basaloid morphology and variable amounts of stroma to the MSRSGC categories.

Of greatest importance for patient management is the accurate separation of benign from malignant neoplasms. The MSRSGC addresses this issue by having four categories most useful for assignment of salivary gland neoplasms. The benign and malignant categories indicate a high level of confidence by the cytopathologist that they classify a given neoplasm as either definitely benign or definitively malignant. Because the authors of the MSRSGC recognized that the cytomorphologic appearance of some specimens is not clearly benign or malignant, two indeterminate categories (SUMP and suspicious for malignancy) were developed to maintain a high diagnostic accuracy for the benign and malignant categories. The operational characteristics of the MSRSGC categories have been well established for the evaluation of salivary gland neoplasms in general but less is known about the diagnostic accuracy of the MSRSGC and its operational characteristics for the diagnostically difficult group of basaloid salivary gland neoplasms.

We accordingly investigated the accuracy of the MSRSGC for basaloid neoplasms by calculating malignancy risk for each of the four diagnostic categories most associated with categorization of neoplasms. The specific malignancy risks for each of the four categories were compared for basaloid neoplasms and salivary gland neoplasms in general (Table 3). As would be expected for a diagnostically difficult set of neoplasms, malignancy risk was higher in the categories neoplasm benign, SUMP and suspicious for malignancy for the basaloid neoplasms than for neoplasms in general. This supports the hypothesis that basaloid neoplasms are more difficult to classify as benign or malignant than salivary gland neoplasms as a whole. Calculated sensitivity and specificity for basaloid neoplasm classification were 90% and 91% respectively. Published data for the MSRSGC analyzing unselected populations of salivary gland lesions show sensitivities varying from 72% to 95%.[18-20] The specificity ranged from 78% to 100%[18-20] These ranges for sensitivity and specificity overlap suggesting that the MSRSGC when used for classifying basaloid neoplasms is as accurate as when classifying salivary gland neoplasms in general. Predictive values of a negative test and of a positive test were also high being 92% and 90% respectively. Data from these studies suggest that the indeterminant category of SUMP aids in maintaining the high diagnostic accuracy of the MSRSGC even for basaloid neoplasms by placing especially diagnostically difficult lesions in this indeterminate category[18-20].

The consensus categories of neoplasms benign and malignant demonstrated good diagnostic accuracy but no correlation existed between neoplasm type and the category SUMP ($p=0.72$). Moreover, there was no significant association between the category SUMP and surgical pathology diagnoses when classified as benign or malignant ($p=0.64$). This suggests that the SUMP category is used by cytopathologists when they are completely unsure if a specimen is benign or malignant. This aids in keeping the diagnostic accuracy for the definitive categories neoplasm benign and malignant high.

A prior study has demonstrated a chance corrected agreement of 0.42 and a Cohen's kappa of 0.71 indicating a substantial agreement among observers for category assignment of a large series of

unselected salivary gland fine-needle aspirates (FNAs).[21] Another study has reported similar results for interobserver agreement in assignment of salivary gland FNAs to Milan System categories.[19] That study demonstrated a Fleiss' kappa for overall categorization agreement of 0.69.[22] We thought that interobserver agreement might be poorer for the diagnostically difficult category of basaloid neoplasms. This hypothesis is supported by the study of Lubin, et al[22] where 33 cases classified as basaloid neoplasms with variable types of stroma disclosed Fleiss' kappas varying from 0.59 to 0.11. Our study documented a Cohen's kappa of 0.27 supporting our hypothesis and confirming the finding of Lubin, et al[22] that diagnosis of basaloid neoplasms is difficult and associated with only fair agreement between observers.

To clarify how the MSRSGC classified basaloid neoplasms, we examined the accuracy of category assignment according to both individual observer assigned categories and consensus categories based on final surgical pathologic diagnosis. The average agreement of individual reviews for the category benign neoplasm was 90.3% (kappa =0.49) and for a malignant categorization it was 92.1% (kappa = 0.58) but average agreement for the SUMP category was 88.8% (kappa = 0.28). Thus agreement for categorization of a specimen as SUMP was poorer than for either the benign or malignant categories. Other studies[23] have also found poor interobserver agreement for the SUMP category. These findings suggest that a SUMP designation for a specimen was given when an observer was unsure as to whether a specimen was benign or malignant and SUMP represented a category of "last resort". This supports the utility of SUMP category despite its poor interobserver reproducibility. When individual categorizations were grouped as benign (atypia of uncertain significance, non-neoplastic, and benign neoplasm) and malignant, sensitivity was 76% and specificity was 77% but when consensus categorizations were used, sensitivity was 92% and specificity was 100%. Thus consensus categorizations were superior in predicting a final benign or malignant diagnosis (Table 9). However, we found no significant correlation between consensus categories and final histologic diagnosis (p=0.27). Consensus

categorization was fairly evenly divided between the categories benign neoplasm (26.7%), malignant (22%) and SUMP (20%). Because of the superior diagnostic accuracy of consensus categorization, we used the consensus categories for further data analysis.

Interobserver agreement for categorization of samples of a basaloid morphology with varying amounts of stroma is fair with a Fleiss' kappa varying from 0.11 to 0.59 depending on nuclear morphology, type of stroma and study reporting the results.[22] Approximately 20% of cases with a basaloid morphology are placed in the SUMP category but interobserver agreement for this category is only fair with a Cohen's kappa of 0.28 in our study and a reported Fleiss' kappa of 0.024.[23] These findings suggest that while the SUMP category may be clinically useful, it is a category with only fair to poor reproducibility and ability to predict the type of neoplasm present or the benign or malignant behavior of neoplasms with a basaloid morphology and variable amounts of stroma. This is particularly important since neoplasms with a basaloid morphology are very difficult to diagnose cytologically but the presence of the SUMP category in the MSRSGC maintains the accuracy of the benign and malignant categories. The interobserver agreement associated with the MSRSGC for all salivary gland FNA samples is superior (kappa = 0.42) to that achieved when only basaloid neoplasms are analyzed (kappa =0.2). These results confirm the difficulty in categorizing basaloid neoplasms. The use of the SUMP and suspicious for malignancy categories helps maintain the high diagnostic accuracies for the benign and malignant categories.

References

1. Layfield LJ, Tan P, Glasgow BJ. Fine-needle aspiration of salivary gland lesions. Comparison with frozen section and histologic findings. Arch Pathol. Lab Med 1987; 111(4): 346-53.
2. Chan MK, McGuire LJ, King W, Li AK, Lee AC. Cytodiagnosis of 112 salivary gland lesions correlation with histologic and frozen section diagnosis. Acta Cytol. 1992; 36(3): 353-63.
3. Cardillo MR. Salivary gland masses: the diagnostic value of fine-needle aspiration cytology. Arch Anat Cytol Pathol. 1990; 38(1-2): 26-32.
4. Schindler S, Nayar R, Dutra J, Bedrossian, CW. Diagnostic challenges in aspiration cytology of the salivary glands. Semin Diagn Pathol. 2001; 18(2): 124-46.
5. Pusztaszen M, Braunschweig R, Mihaescu A. Pleomorphic adenoma with predominant plasmacytoid myoepithelial cells: a diagnostic pitfall in aspiration cytology. Case report and review of the literature. Diagn Cytopathol. 2009; 37(1): 56-60.
6. Chrabanska M, Kiczmer P, Drozdzowska B. Salivary gland lesions: diagnostic reliability and challenges of fine needle aspiration cytology. Int J Clin Exp Pathol. 2021; 14(1): 54-62.
7. Layfield LJ. Fine needle aspiration cytology of a trabecular adenoma of the parotoid gland. Acta Cytol. 1985; 29(6): 999-1002.

8. Kapadia SB, Dusenbery D, Dekker A. Fine needle aspiration of pleomorphic adenoma and adenoid cystic carcinoma of salivary gland origin. Acta Cytol. 1997; 41(2): 487-92.
9. Stanley MW, Horwitz CA, Rollins SD, Powers CN, Bardales RH, Korourain S, Stern SJ. Basal cell (monomorphic) and minimally pleomorphic adenomas of the salivary glands. Distinction from the solid (anaplastic) type of adenoid cystic carcinoma in fine-needle aspiration. Am J Clin Pathol. 1996; 106(1): 35-41.
10. Jurczyk M, Peevey JF, Vande Harr MA, Lin X. Pitfalls of fine-needle aspiration cytology of parotid membranous basal cell adenoma. A review of pitfalls in FNA cytology of salivary gland neoplasms with basaloid cell features. Diag Cytopathol. 2015; 43(5): 437-7
11. Geisinger KR, Weidner. Aspiration cytology of salivary glands. Semin Diagn Pathol. 1986; 3(3): 219-26.
12. Verma K, Kapila K. Role of fine needle aspiration cytology in diagnosis of pleomorphic adenomas. Cytopathol. 2002; 13(2): 121-7.
13. Hughes JH, Volk EE, Wilbar DC. Cytopathology Resource Committee. College of American Pathologists. Pitfalls in salivary gland fine-needle aspiration cytology: lessons from the College of American Pathologists Interlaboratory Comparison Program in Non-gynecologic Cytology. Arch Pathol Lab Med. 2005; 129(1): 26-31.
14. Nagel H, Hutze HJ, Laskawi R, Chilla R, Droese M. Cytologic diagnosis of adenoid cystic carcinoma of salivary glands. Diagn Cytopathol. 1999; 20(6): 358-66.
15. Hara H, Oyama T, Suda K. New criteria for cytologic diagnosis of adenoid cystic carcinoma. Acta Cytol. 2005; 49: 43-50.
16. Cantley RL. Fine-needle aspiration cytology of cellular basaloid neoplasms of the salivary gland. Arch Pathol Lab Med. 2019; 143(11): 1338-1345.

17. Baloch Z, Field AS, Katabi N, Wenig BM. The Milan System for Reporting Salivary Gland Cytopathology. In: Faquin WC, Rossi ED editors. The Milan System for Reporting Salivary Gland Cytopathology 1st ed. Cham Switzerland: Springer. 2018. 1-9.
18. Reerds STH, Van Engen-VanGransven ACH, vanden Hoogen FJA, Takes RP, Marres HAM, Honings J. Validation of the Milan System for Reporting Salivary Gland Cytopathology and the diagnostic accuracy of FNA cytology for submandibular gland lesions. Cancer Cytopathol. 2021; Nov 30, doi: 10.1002/cncy22532.
19. Rossi ED. The Milan System, from its introduction to its current adoption in the diagnosis of salivary gland cytology. J Mol Pathol. 2021; 2(2): 114-122.
20. Sadullahoglu C, Yildirim HT, Nergiz D, et al. The risk of malignancy according to Milan reporting system of salivary gland fine-needle aspiration with Beeton Dickinson SurePath liquid-based processing. Diagn Cytopathol. 2019; 47(9): 863-868.
21. Layfield LJ, Esebua M, Yang Z, Vazmitsel M, Giorgadze T, Schmidt R. The Milan system for reporting salivary gland cytopathology: A study of interobserver reproducibility. Diagn Cytopathol. 2019; 47(8): 765-68.
22. Lubin DJ, Griffith CC, Buonocore DJ, Wei XJ, Lin O. Putting morphology to the test: An established classification scheme reliably stratifies salivary gland cytology by risk of malignancy with substantial interobserver agreement. Cancer Cytopathol. 2021; 129(4): 291-303.
23. Hollyfield JM, O'Connor SM, Maygarden SJ, Greene KG, Scanga LR, Tang S, et al. Northern Italy in the American South: Assessing interobserver reliability within the Milan System for Reporting Salivary Gland Cytopathology. Cancer Cytopathol. 2018; 126(6): 390-396.

Table 1: Distribution of tumor types established by histopathologic examination

Tumor Type	Frequency	Percent
ACC	19	42.2
BCC	1	2.2
BMT	10	2.2
MA	13	28.9
PLGA	2	4.4
Total	45	100

ACC = Adenoid cystic carcinoma, BCC = Basaloid cell carcinoma,
BMT = Benign mixed tumor, MA = Monomorphic adenoma,
PLGA = Polymorphous low-grade adenocarcinoma

Table 2: Reviewer categories for all case and final surgical pathology diagnosis

		Reviewer #							
Case #		1	2	3	4	5	6	7	Tumor Type
1		BN	SUMP	BN	SUMP	BN	SUMP	SUMP	MA
2		BN	BN	BN	SUMP	SUMP	SUMP	SUMP	ACC
3		M	BN	SUMP	SUMP	SUMP	SUMP	BN	ACC
4		BN	BN	BN	BN	BN	BN	BN	BMT
5		M	SM	M	M	M	M	M	ACC
6		M	M	M	M	M	M	M	ACC
7		SUMP	SM	M	SUMP	M	SUMP	SUMP	MA
8		BN	BN	BN	BN	SUMP	BN	BN	MA
9		BN	SM	SM	M	SUMP	SUMP	SM	ACC
10		BN	SM	BN	SUMP	SUMP	BN	BN	BMT
11		BN	BN	BN	BN	SUMP	SUMP	BN	BMT
12		M	SM	BN	BN	BN	BN	BN	MA
13		BN	M	SUMP	M	SUMP	SM	BN	Basal Cell CA
14		SUMP	AUS	SUMP	AUS	AUS	SUMP	BN	MA
15		M	M	M	M	M	M	M	ACC
16		BN	SUMP	M	SUMP	SM	M	M	ACC
17		BN	SUMP	SUMP	M	SUMP	SUMP	BN	MA
18		BN	BN	SUMP	BN	BN	BN	BN	MA
19		SUMP	SUMP	SM	SUMP	SM	SUMP	SUMP	ACC
20		SUMP	M	BN	SM	SUMP	SUMP	M	MA
21		SUMP	BN	AUS	BN	AUS	SUMP	BN	ACC
22		BN	SM	SUMP	M	SUMP	SM	SUMP	MA
23		M	BN	SM	SUMP	SM	SM	M	PLGA
24		SM	BN	BN	SUMP	SUMP	SUMP	BN	MA
25		M	M	M	M	M	SUMP	M	ACC
26		BN	SUMP	BN	SM	SUMP	M	SUMP	BMT
27		BN	BN	BN	BN	BN	BN	BN	ACC
28		BN	BN	BN	BN	SUMP	SUMP	BN	ACC
29		BN	BN	SUMP	SUMP	SUMP	BN	SM	ACC
30		M	M	M	M	M	SM	M	BMT
31		M	SUMP	M	SUMP	M	M	BN	ACC
32		SUMP	BN	SM	SUMP	SUMP	SM	BN	MA
33		SM	SUMP	SUMP	SUMP	SUMP	SM	BN	BMT
34		SUMP	BN	BN	SUMP	SUMP	SM	SUMP	BMT
35		BN	BN	BN	SUMP	SUMP	SUMP	BN	ACC
36		SUMP	SUMP	BN	BN	SUMP	SUMP	BN	PLGA
37		M	M	M	M	M	M	M	ACC
38		BN	SUMP	M	M	M	M	SUMP	MA
39		SM	AUS	SM	NN	SM	SM	SUMP	ACC
40		BN	SM	SUMP	SM	SUMP	SM	SUMP	MA
41		M	M	M	M	M	M	M	ACC
42		BN	BN	BN	SM	BN	SUMP	BN	BMT
43		SUMP	M	M	AUS	M	SUMP	BN	BMT
44		BN	BN	M	SM	SUMP	BN	BN	ACC
45		BN	NN	BN	AUS	BN	AUS	BN	BMT

ND = Non-diagnostic

NN = Non-neoplastic

AUS = Atypia

SUMP = Salivary Gland Neoplasms of Uncertain Malignant Potential

BMT = Benign Mixed Tumor

PLGA = Polymorphous Low-Grade Adeno CA

SM = Suspicious for Malignancy

M = Malignant

BN = Benign Neoplasm

MA = Monomorphic Adenoma

ACC = Adenoid Cystic CA

Basal Cell CA = Basal Cell Carcinoma

Table 3: Malignancy risks for basaloid neoplasms by Milan System categories compared to Milan System categories reported in the literature (unselected neoplasms).

Category	Basaloid Neoplasms[10,16]	All Neoplasms[1-6,20,22]
Neoplasm -Benign	8%	<5%
SUMP	44%	35%
Suspicious for Malignancy	100%	60%
Malignant	90%	90%
Combined Suspicious and Malignant	91%	unknown

Table 4: Impact of experience on diagnostic accuracy

Reviewer	Years Exp	Accuracy > Reviewer	Accuracy Group
3	<10	80.0	75.6
6	<10	71.1	
4	10-19	64.4	71.1
5	10-19	77.8	
1	>20	75.6	69.6
2	>20	57.8	
7	>20	75.6	

Table 5: Consensus Diagnosis Category Assignment

Consensus Diagnosis	Frequency	Percent
None	13	28.9
BN	12	26.7
SUMP	9	20.0
SM	1	2.2
M	10	22.2
Total	45	100.0

BN = Benign neoplasm, SUMP = Salivary gland neoplasm of unknown Malignant potential, SM = Suspicious for malignancy, M = Malignant

Table 6: Cross tabulation of Histologic Diagnosis with SUMP category

Tumor Type	Consensus Dx		Total
	Other	SUMP	
ACC	16	3	19
BCC	1	0	1
BMT	8	2	10
MA	10	3	13
PLGA	1	1	2
Total	36	9	45

Table 7: Cross tabulation of Category of Final (correct) Diagnosis and SUMP

Final Dx	SUMP Consensus Dx		Total
	Other	SUMP	
Benign	19	5	24
Malignant	17	4	21
Total	36	9	45

Table 8: Cross tabulation of Final Histologic Diagnosis with Consensus Diagnosis

Tumor Type	Concensus Diagnosis				Total
	No Consensus	BN	SUMP	M	
ACC	3	4	3	9	19
BCC	1	0	0	0	1
BMT	2	5	2	1	10
MA	6	3	3	1	13
PLGA	1	0	1	0	2
Total	13	12	9	11	45

Table 9: The table shows the average agreement for 15 reviewer pairs

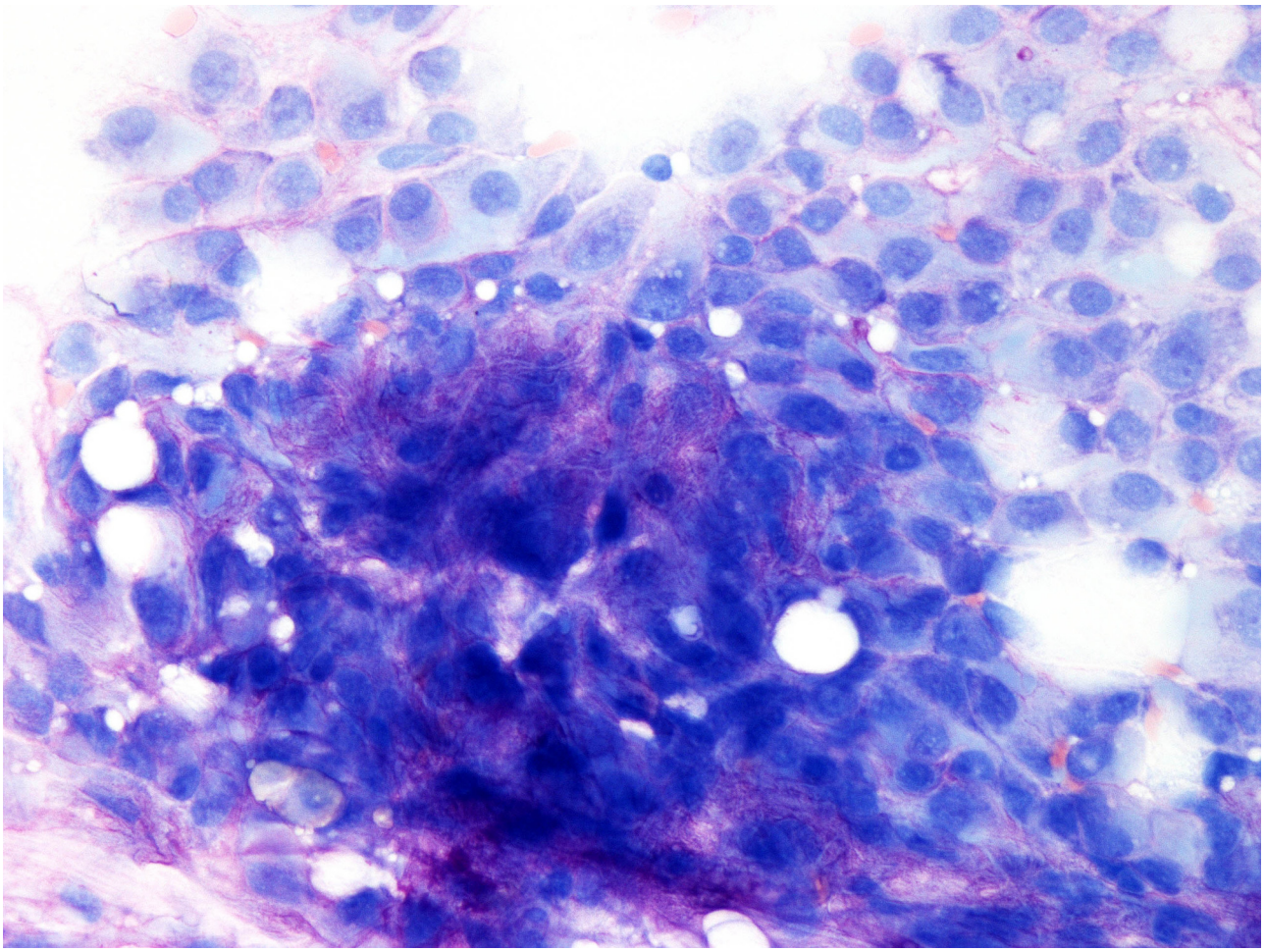
Cytology Diagnosis	Observed Agreement (95% C1, p value vs SUMP)	Kappa (95% C1, p value vs SUMP)
Sump	88.8% [87.8 -89-7]	0.28 [0.23 -0.33]
Benign (BN or AUS)	90.3% [89.3 - 91.5%, 0.02]	0.49 [0.45 - 0.53, 0.01]
Malignant (SM or M)	92.1% [91.2 - 93.0, <0.0005]	0.58 [0.53 - 0.62, 0.001]

Table 10: Diagnostic Accuracy of Consensus Diagnosis compared with histologic classification. The sensitivity was 92% (95% CI: 62-100) and the specificity was 100% (95% CI: 72-100).

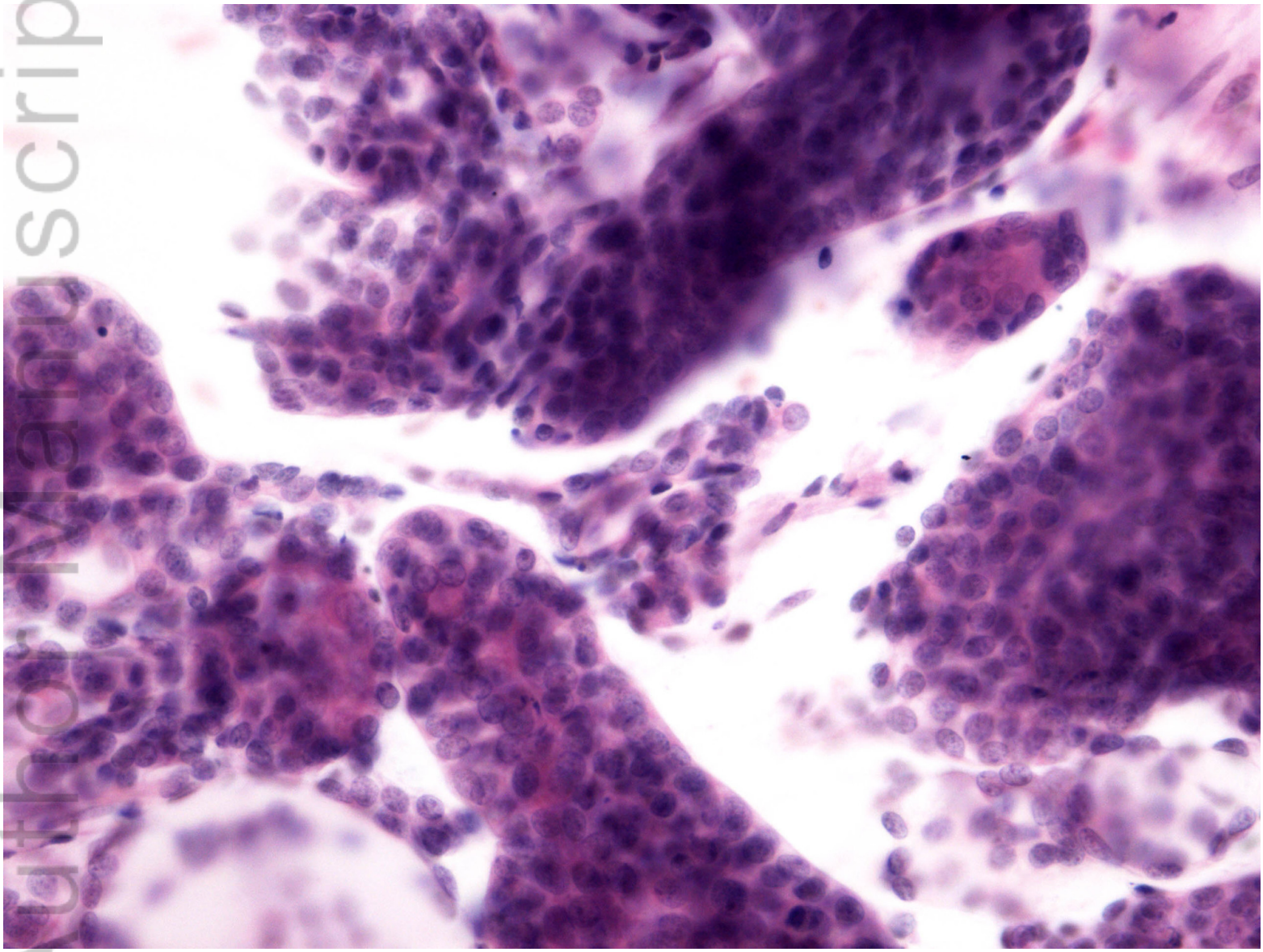
Consensus Diagnosis	Final Diagnosis		Total
	BN	M	
BN	11	1	12
M	0	11	11
None	8	5	13
SUMP	5	4	9
Total	25	11	45

Table 11: Diagnostic Accuracy of Definitive Individual Diagnosis compared with final histologic classification (exclude SUMP). The sensitivity of individual diagnoses was 76% (95% CI: 67-84) and the specificity was 77% (95% CI: 68 -84).

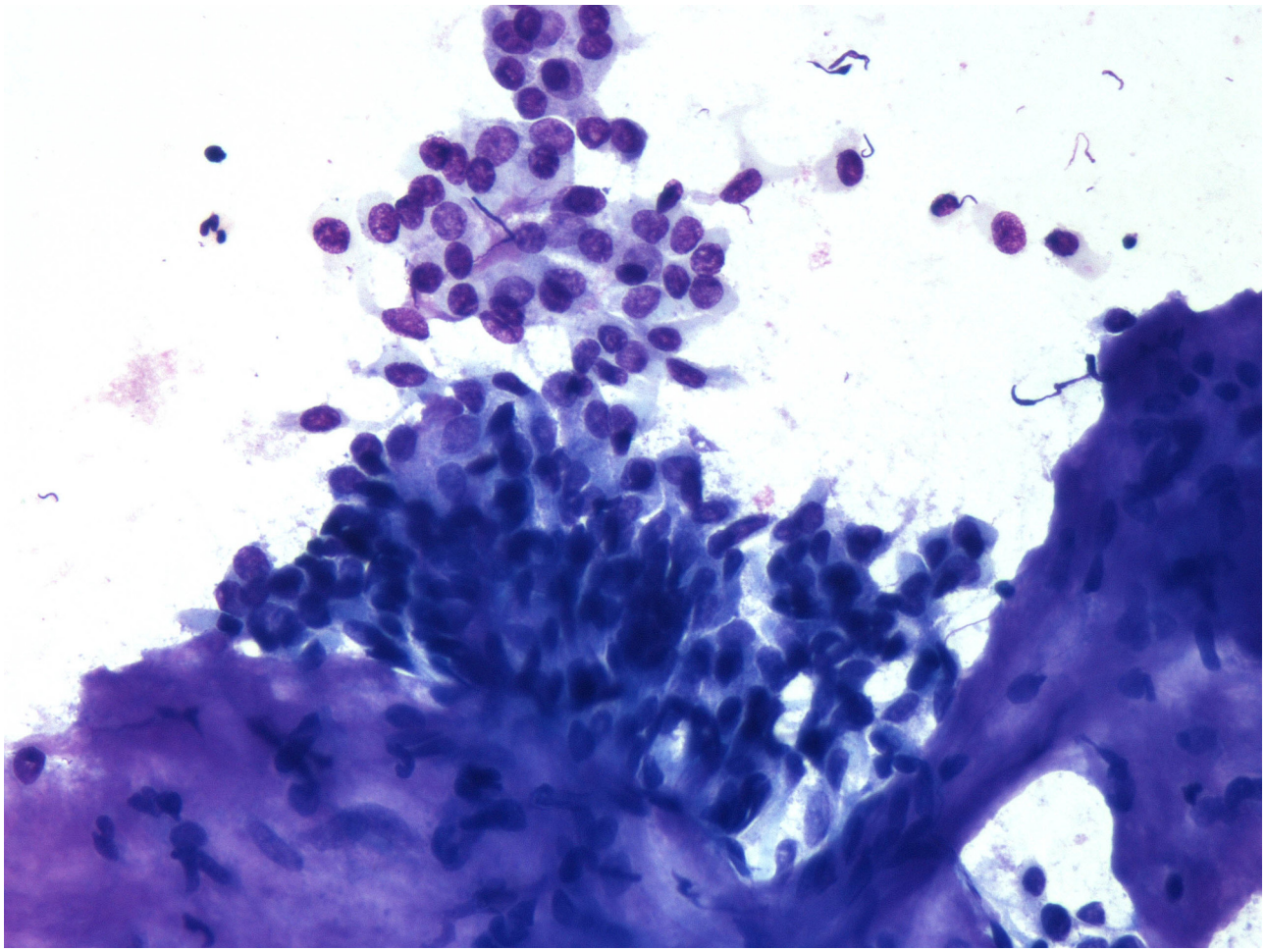
Individual Diagnosis	Final Diagnosis		Total
	BN	M	
NN, BN, AUS	87	26	113
SM, M	26	84	110
Total	113	110	223



DC_24962_Figure 1 -Basaloid 3-25-22.jpg



DC_24962_Figure 2- Basaloid -3-25-22.jpg



DC_24962_Figure 3 -Basaloid - 3-25-22.jpg

Legends:

Figure 1: Photo micrograph of case 42 for there was a high degree of agreement that the neoplasm was benign. The smears show many plasmacytoid myoepithelial cells lying in a back ground composed of scant myxoid to fibrillar stroma (Diff Quik, X600).

Figure 2: Photomicrograph of case 28 which was associated with a high degree of agreement that the neoplasm was malignant. Smears show clusters of small oval cells. The clusters often have a “finger in glove” configuration characteristic of adenoid cystic carcinoma (H+E. X400).

Figure 3: Photomicrograph of case 40 which had a low degree of interobserver agreement. The smear contained irregular clusters and sheets of small basaloid to short spindle cells. The cells often exfoliated off the larger cell groups. The neoplasm is a monomorphic adenoma (Diff Quik, X400).