Original Article

Accuracy of Parotid Gland FNA Cytology and Reliability of the Milan System for Reporting Salivary Gland Cytopathology in Clinical Practice

Sam T. H. Reerds, MD [iD] [1]; Adriana C. H. Van Engen–Van Grunsven, MD, PhD[2];
Frank J. A. van den Hoogen, MD, PhD[1]; Robert P. Takes, MD, PhD [iD] [1]; Henri A. M. Marres, MD, PhD[1];
and Jimmie Honings, MD, PhD [iD] [1]

BACKGROUND: Differentiating between malignant and benign salivary gland tumors with fine-needle aspiration cytology (FNAC) can be challenging. This study was aimed at testing the validity of the Milan System for Reporting Salivary Gland Cytopathology (MSRSGC) and at assessing possible differences in the sensitivity and specificity of parotid gland FNAC between dedicated head and neck (H&N) centers, subdivided into head and neck oncology centers (HNOCs) and head and neck oncology affiliated centers (HNOACs), and general hospitals (GHs). **METHODS:** The Dutch Pathology Registry (PALGA) database was searched for patients who had undergone a salivary gland resection between January 1, 2006, and January 1, 2017, and had a preoperative FNAC result. The FNAC reports were retrospectively assigned to MSRSGC categories. The risk of malignancy (ROM) was calculated for each category. The sensitivity and specificity for diagnosing malignancy were calculated and compared among HNOCs, HNOACs, and GHs. **RESULTS:** In all, 12,898 FNAC aspirates were evaluated. The ROMs for each category were as follows: 12.5% in MSRSGC I, 10.3% in MSRSGC II, 29% in MSRSGC III, 2.3% in MSRSGC IVa, 28.6% in MSRSGC IVb, 83% in MSRSGC V, and 99.3% in MSRSGC VI. The sensitivity of FNAC was highest in HNOCs (88.1%), HNOACs scored lower (79.7%), and GHs had a sensitivity of 75.0%. **CONCLUSIONS:** The MSRSGC is a valid tool for reporting parotid gland FNAC; therefore, these results strongly advocate its use. On the basis of the higher sensitivity of FNAC in dedicated H&N centers, the authors recommend that GHs use the presented management strategies to help to minimize the chances of a preoperative misdiagnosis. *Cancer Cytopathol* 2021;129:719-728. © 2021 The Authors. *Cancer Cytopathology* published by Wiley Periodicals LLC on behalf of American Cancer Society This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

KEY WORDS: cytological techniques; parotid gland; parotid neoplasms; retrospective studies.

INTRODUCTION

A mass in the parotid gland is potentially neoplastic. The diagnosis of lesions in the parotid can be challenging because there are more than 40 different benign and malignant salivary gland tumors. In addition, various types of metastatic tumors can be found in the intraparotid lymph nodes.[1] Treatment of both benign and malignant tumors mostly relies on surgical resection. Furthermore, a parotid gland mass can be of nonneoplastic origin. Sialadenosis, sialolithiasis, sialadenitis, oncocytosis, sialometaplasia, cysts, and reactive enlarged lymph nodes can all occur in the parotid gland and cause a mass.

The primary diagnostic technique used for the preoperative assessment of a parotid gland mass is fine-needle aspiration cytology (FNAC). The assessment of a mass using FNAC can help the clinician to decide whether surgical intervention is necessary (is it of neoplastic or nonneoplastic origin?). Furthermore, it can play an essential role in determining the timing (is it a benign or malignant neoplasm?) and extent of surgical treatment (is the tumor subtype of a high histopathological grade or low grade?).

Unfortunately, FNAC is known to have its limitations. Although some specific tumor types are mostly easily diagnosed, the differentiation between certain types of salivary gland tumors or the distinction between a benign tumor and a malignant tumor can prove challenging. This is, among other things, caused by the overlap in cell types between different salivary gland tumor types and the fact that additional staining is often of little help. On account of the technique, it is also impossible to determine signs of malignancy such as invasive growth and perineural or vaso-invasive growth. Moreover, there is the possibility of a sampling error.

A previous systematic review performed by Liu et al[2] showed that parotid gland FNAC had an overall sensitivity of 78% and an overall specificity of 97.8% for correctly identifying the tumor's dignity (eg, benign, malignant or non-neoplastic). In the same review, a subgroup analysis was performed for FNAC under ultrasound guidance. This showed a sensitivity of 84.8% and a specificity of 98%. Moreover, 13.3% of all FNACs performed showed indeterminate or nondiagnostic results. As a result of this sensitivity, there is a significant risk for false-negative results (in other words, malignant tumors falsely diagnosed as benign tumors or nonneoplastic lesions). Furthermore, indeterminate and nondiagnostic FNAC diagnoses lack clarity for the patient and the treating clinician.

To provide a more objective and reproducible measure for clinicians, an international group of experts supported by the American Society of Cytopathology and the International Academy of Cytology developed a categorical system in 2018 called the Milan System for Reporting Salivary Gland Cytopathology (MSRSGC).[3,4] The MSRSGC contains different diagnostic categories in which FNAC results are subdivided. The American Society of Cytopathology also estimated the risk of

TABLE 1. Milan System for Reporting Salivary Gland Cytopathology Categories, Appurtenant Risk of Malignancy, and Advised Management Strategy as Reported by Faquin et al[4]

Diagnostic Category	Risk of Malignancy, %	Management
I. Nondiagnostic	25	Clinical and radiological correlation/repeat FNAC
II. Nonneoplastic	10	Clinical follow-up and radiological correlation
III. AUS	20	Repeat FNAC or surgery
IVa. Neoplasm: benign	<5	Surgery or clinical follow-up
IVb. SUMP	35	Surgery
V. Suspected malignant	60	Surgery
VI. Malignant	90	Surgery

Abbreviations: AUS, atypia of unknown significance; FNAC, fine-needle aspiration cytology; SUMP, salivary gland neoplasm of unknown malignant potential.

malignancy (ROM) within each group and subsequently provided an advised management strategy (Table 1).

Numerous studies have assessed the accuracy of cytopathological evaluation using the MSRSGC tool; these have concluded that this tool can make a valuable contribution to the management of a parotid gland mass.[5-14] Unfortunately, most of these studies have lacked large groups and have been mostly single-institution or bi-institutional studies. As a result, the claim has been made that large-scale, multicenter studies are imperative for testing the reliability and validity of the MSRSGC classification.[15]

Head and neck (H&N) oncology care in the Netherlands is centralized in 14 hospitals: 8 head and neck oncology centers (HNOCs) and 6 head and neck oncology affiliated centers (HNOACs). HNOCs are tertiary, mostly academic referral centers, whereas HNOACs are general hospitals (GHs) that are closely affiliated with them. The HNOACs use the same treatment protocols as the related HNOCs. No previous studies have compared the diagnostic accuracy of FNAC between hospitals (ie, HNOCs, HNOACs, and GHs).

Our primary objective was to study the diagnostic accuracy of FNAC and to test the validity of using the MSRSGC classification in a large nationwide cohort. The secondary objective was to examine the differences in diagnostic accuracy for parotid gland FNAC among HNOCs (n = 8), HNOACs (n = 6), and GHs (n = 36). Accordingly, we contemplate management strategies for GHs based on the possible differences between dedicated H&N centers and GHs.

MATERIALS AND METHODS

Patients

The PALGA database, "the nationwide network and registry of histo- and cytopathology in the Netherlands,"[16] was consulted to search for patients who had undergone a salivary gland resection between January 1, 2006, and January 1, 2017. Information on age, the date of examination, the side of the lesion, and the type of hospital where the diagnosis was rendered (HNOC, HNOAC, or GH) was included in the database.

Study Approval

This study was approved by the scientific and privacy committee of PALGA. Because of the anonymous patient data collection and its retrospective nature, the study did not fall within the remit of the Medical Research Involving Human Subjects Act.

Inclusion and Exclusion

Histopathological diagnoses were classified according to the 2005 World Health Organization classification for salivary gland tumors,[17] which was most appropriate to the search period. Tumors classified as malignant epithelial tumors, borderline tumors, benign epithelial tumors, other epithelial lesions, or soft tissue lesions according to the World Health Organization classification for salivary gland tumors were included in the study. Metastatic tumors to the parotid were included because the MSRSGC classification was designed to also differentiate between these. Lymphomas were excluded because surgery is not the treatment of choice for lymphoma. Therefore, this group was underrepresented in the histopathological resections, and as a result, these FNACs lacked a definitive diagnosis for comparison. Lesions outside the parotid gland (eg, the submandibular or the sublingual gland) were excluded. Patients were excluded if the FNAC was performed more than 1 year before the resection and if there was a histopathological result of the lesion before the first FNAC because of the bias that this may have presented to the pathologist.

Categorization and Analysis

The MSRSGC guideline was used to categorize the cytopathological reports retrospectively by one of the authors (S.T.H.R.). In case of uncertainty, a second reviewer (A.C.H.V.E.V.G.) was consulted.

Each FNAC result was compared with the definitive histopathological diagnosis as the golden standard; it was stated if the result was discordant or concordant with the FNAC result. In case of revision of the histopathological diagnosis, the revised diagnosis was considered the golden standard. The ROM was determined for each MSRSGC category. Consequently, the sensitivity and specificity for diagnosing malignancy were calculated as measures of diagnostic accuracy. For this analysis, the suspected malignant (MSRSGC V) and malignant (MSRSGC VI) groups were classified as positive cytopathological tests, and the nonneoplastic (MSRSGC II) and benign salivary gland neoplasm (MSRSGC IVa) groups were categorized as negative test results. Indeterminate (MSRSGC III or IVb) and nondiagnostic (MSRSGC I) results were excluded from the analysis of sensitivity and specificity; these results were separately reported. Cytopathological revisions were excluded from the diagnostic accuracy analysis because of the possible bias that these may present.

Statistical Analysis

Sensitivity and specificity were calculated along with their respective 95% confidence intervals. χ^2 tests were used to test for differences between sensitivity and specificity between the different types of hospitals. Two-sided P values < .05 were considered statistically significant. Statistical analysis was performed with SPSS version 26.0 (IBM Corp, Armonk, New York).

RESULTS

In total, 24,164 patients with a salivary gland resection were gathered from the PALGA database. The inclusion and exclusion process is summarized in Figure 1.

During the study period, 12,898 FNAC aspirates from the parotid gland were taken from a total of 9672 patients who underwent subsequent resection of the lesion within 1 year after the FNAC. The study group consisted of 4807 male patients (49.7%) and 4865 female patients (50.3%). The mean age at first FNAC was 54.8 years (range, 0-98 years; SD, 15.5 years). The left and right parotid glands were equally involved (50% and 49.7%). In 0.3% of cases, the side was unknown. A small portion of patients (0.9%) had a bilateral mass surgically removed.

Twenty-nine percent of all FNAC aspirates were evaluated in an HNOC, 19.3% were evaluated in an HNOAC, and 51.6% were evaluated in a GH. In 81.8% of the
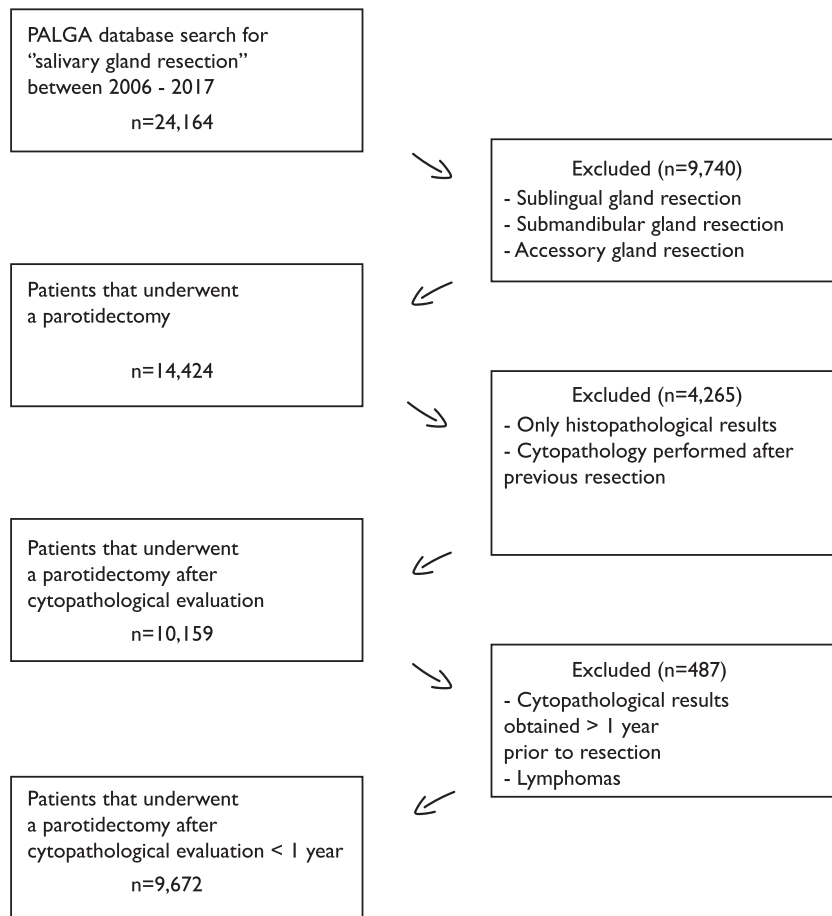
Figure 1. Inclusion and exclusion flow diagram. PALGA indicates Dutch Pathology Registry.

TABLE 2. Distribution of Cytopathological Results According to the Milan System for Reporting Salivary Gland Cytopathology and Corresponding Risks of Malignancy After Correlation With Histopathological Results

Diagnostic Category	Distribution, %	Risk of Malignancy, %
I. Nondiagnostic	19.0	12.5
II. Nonneoplastic	2.2	10.3
III. AUS	3.2	29
IVa. Benign	61.4	2.3
IVb. SUMP	6.4	28.6
V. Suspected malignant	3.0	83
VI. Malignant	4.7	99.3

Abbreviations: AUS, atypia of unknown significance; SUMP, salivary gland neoplasm of unknown malignant potential.

patients, only 1 FNAC was performed. A second FNAC was performed in 15%, and a third FNAC was performed in 2.8%. Only 0.4% of patients had 4 or more aspirates taken (up to 7 FNACs). After resection, the final histopathological diagnosis yielded a malignant neoplasm in 12.4%, a benign tumor in 84.2%, and nonneoplastic disease in 3.4%.

The overall distribution of FNAC aspirates classified according to the MSRSGC is provided in Table 2. The most frequent preoperative FNAC result was a benign salivary gland tumor (61.4%), which was followed by a nondiagnostic result (19.0%), a salivary gland neoplasm of unknown malignant potential (SUMP; 6.4%), and a malignant tumor (4.7%). The same table also shows the ROM for each category. These were high in both the malignant (99.3%) and suspected malignant categories (83%). The SUMP and atypia of unknown significance (AUS) categories had ROMs of 28.6% and 29%, respectively. The ROMs for the nondiagnostic category (12.5%), the nonneoplastic category (10.3%), and benign salivary gland tumors (2.3%) were lower.

The majority of all cytopathological smears (57%) were taken and evaluated at a GH. The HNOCs and HNOACs evaluated 21.7% and 21.3%, respectively.

The sensitivity of FNAC was highest in the dedicated H&N centers: HNOCs had a sensitivity of 88.1%,

TABLE 3. Summary of the Sensitivity and Specificity With Their Respective 95% CIs and Nondiagnostic or Indeterminate Diagnoses for the Different Types of Hospitals

Type of Hospital	Total, No.	%	Sensitivity, % (95% CI)	Specificity, % (95% CI)	Nondiagnostic/Indeterminate, %		
					MSRSGC I	MSRSGC III	MSRSGC IVb
HNOC	2527	21.7	88.1 (84.7-90.9)	98.4 (97.6-98.9)	19.4	3.4	8.9
HNOAC	2486	21.3	79.7 (73.4-84.8)	99.2 (98.7-99.6)	19.2	3.4	6.4
GH	6644	57.0	75.0 (70.6-78.9)	99.3 (99.0-99.5)	18.8	3.0	5.4
Total	11,657	100	81.2 (78.7-83.6)	99.1 (98.8-99.3)	19	3.2	6.4

Abbreviations: CI, confidence interval; GH, general hospital; HNOAC, head and neck oncology affiliated center; HNOC, head and neck oncology center; MSRSGC, Milan System for Reporting Salivary Gland Cytopathology.

HNOACs had a lower sensitivity of 79.7%, and GHs had a sensitivity of 75%.

The sensitivity and specificity of FNAC differed significantly between HNOCs and HNOACs ($P = .006$ and $P = .034$, respectively) and between HNOCs and GHs ($P < .001$ and $P = .002$, respectively). When we compared HNOACs and GHs, there were no differences in sensitivity or specificity found ($P = .205$ and $P = .803$, respectively). Table 3 summarizes these results.

Most benign tumors and nonneoplastic parotid gland lesions were resected at a GH (Table 4). Malignant tumors, on the other hand, were mostly resected at an HNOC; 15.4% of malignant tumors (n = 185) were removed at a GH.

Among the 185 malignant tumors removed at a GH, 45 tumors (24.3%) had a preoperative diagnosis of a malignant parotid gland tumor or a suspected malignant parotid gland tumor (MSRSGC V or VI), 33 tumors (17.8%) had a preoperative diagnosis of SUMP (MSRSGC IVb), and 9 tumors (4.9%) had a preoperative diagnosis of AUS (MSRSGC III). The other 98 tumors (53.0%) had a nondiagnostic, nonneoplastic, or benign preoperative FNAC result.

FNAC provided a false-negative diagnosis for some of the resected malignant parotid tumors (MSRSGC II/ MSRSGC IVa). The histopathological subtypes associated with the highest false-negative rates on FNAC before surgery are shown in Table 5. Histopathologically proven myoepithelial carcinomas were shown to have the highest false-negative rate (57.1%) on preoperative FNAC.

The exact cytopathological subtype diagnoses within MSRSGC class IVa that proved malignant on histopathological examination (false-negative or false-benign results) were basal cell adenoma (50%, n = 7), myoepithelioma (33.3%, n = 4), and oncocytoma (9.1%, n = 2). More frequently occurring cytopathological diagnoses such as Warthin tumor and pleomorphic adenoma had

TABLE 4. Number of Resections Performed in Each Type of Hospital With Categorization by Tumor Dignity

Type of Hospital	Benign		Nonneoplastic		Malignant	
	No.	%	No.	%	No.	%
HNOC	2683	32.9	130	39.9	861	71.6
HNOAC	1636	20.1	61	18.7	156	13.0
GH	3825	47.0	135	41.4	185	15.4
Total	8144		326		1202	

Abbreviations: GH, general hospital; HNOAC, head and neck oncology affiliated center; HNOC, head and neck oncology center.

TABLE 5. Most Frequent False-Negative Histopathological Diagnoses Arranged by Their Corresponding False-Negative Rates

Type of Malignancy	False-Negatives, No.	False-Negative Rate, %[a]
Myoepithelial carcinoma	12	57.1
Epithelial-myoepithelial carcinoma	19	47.5
Mucoepidermoid carcinoma	42	43.8
Carcinoma ex pleomorphic adenoma	19	32.8
Acinic cell carcinoma	41	29.9
Adenoid cystic carcinoma	11	23.4
Metastatic squamous cell carcinoma	12	5.0

(False negatives on fine-needle aspiration cytology/Total number of histopathological diagnoses) × 100
Only diagnoses with at least 10 false-negative results are included.
[a]The false-negative rates were calculated as follows:

false-negative rates of 2.1% and 1.9%, respectively. The false-negative results in the Warthin tumor group included mucoepidermoid carcinoma (n = 20), acinic cell carcinoma (n = 17), salivary duct carcinoma (n = 4), metastatic squamous cell carcinoma (n = 2), adenocarcinoma not otherwise specified (n = 1), epithelial-myoepithelial carcinoma (n = 1), and carcinoma ex pleomorphic adenoma (n = 1). The false-negative results in the pleomorphic adenoma group included carcinoma ex pleomorphic adenoma (n = 17), epithelial-myoepithelial carcinoma (n = 16),

TABLE 6. MSRSGC Diagnostic Categories and Their Corresponding Risks of Malignancy in Previous Studies

Source	No.	No.[a]	Risk of Malignancy, %							Sensitivity	Specificity
			I	II	III	IVa	IVb	V	VI		
Viswanathan 2018[5]	627	373	6.7	7.1	38.9	5	34.2	92.9	92.3	79.0	98
Savant 2019[6]	199	199	0	0	33	0.8	40.9	100	100	—	—
Dubucs 2019[7]	328	216	34	0	0	3.1	45.5	68.8	100	—	—
Mazzola 2019[8]	375	366	19	11.8	25	5.5	50	71.4	94.6	—	—
Wu 2019[9]	1560	694	18.3	8.9	37.5	2.9	40.7	100	98.3	89	99
Song 2019[10]	893	429	16.1	17.9	30.6	2.2	46.6	78.9	98.5	—	—
Choy 2019[11]	376	376	14.5	26.7	29.3	2.7	19.1	87.5	100	—	—
Chen 2019[12]	1020	349	8.6	15.4	36.8	2.6	32.3	71.4	100	70.4	99.2
Lee 2019[13]	1384	421	10	17.5	29.5	0.5	17.1	83.3	100	76.5	99.1
Mazzola 2020[14]	503	503	19.5	14.3	17.6	3.6	24.6	66.7	96.8	—	—
Total[b]	**6889**	**3550**	**14.1**	**12.7**	**31.6**	**2.7**	**34.8**	**80**	**97.5**	**—**	**—**
Current study	*12,898*	*12,898*	*12.5*	*10.3*	*29*	*2.3*	*28.6*	*83*	*99.3*	*81.2*	*99.1*
MSRSGC classification	*—*	*—*	*25*	*10*	*20*	*<5*	*35*	*60*	*90*	*—*	*—*

Abbreviation: MSRSGC, Milan System for Reporting Salivary Gland Cytopathology.
[a]Histopathologically correlated fine-needle aspiration cytology.
[b]Choy et al's study[11] is not part of the total calculation because of the unavailability of their MSRSGC distribution rates.

acinic cell carcinoma (n = 11), myoepithelial carcinoma (n = 8), adenoid cystic carcinoma (n = 9), adenocarcinoma not otherwise specified (n = 5), salivary duct carcinoma (n = 3), basal cell adenocarcinoma (n = 3), secretory carcinoma (n = 2), polymorphic adenocarcinoma (n = 1), sarcoma (n = 1), large cell carcinoma (n = 1), primary squamous cell carcinoma (n = 1), metastatic squamous cell carcinoma (n = 1), and melanoma (n = 1).

DISCUSSION

FNAC is the standard diagnostic test in the management of salivary gland lesions, but as illustrated previously, it is known to have its limitations. Fortunately, the presented ROMs validate the use of the MSRSGC.

The distribution of the ROMs in our study (Table 2) showed minor differences in comparison with the estimates of the MSRSGC (Table 1). The ROMs of MSRSGC III (AUS), MSRSGC V (suspected malignant), and MSRSGC VI (malignant) proved slightly higher. However, these results still justify the management proposed in the MSRSGC, namely to either repeat FNAC or perform surgery for MSRSGC III and to perform surgery for MSRSGC V and VI.

Previous studies reporting on the ROMs for MSRSGC categories have shown variation in the distribution of ROMs, especially in the MSRSGC I, II, III, and IVb categories.[5-14] These studies are summarized in Table 6. Only studies that included at least 200 patients with histopathological confirmation were included in this summary. The vast majority of these studies also included patients who had no histopathological confirmation, which means that they had only clinical or radiological correlation to verify the assigned MSRSGC category. These patients were mostly represented in the nondiagnostic (MSRSGC I) and nonneoplastic (MSRSGC II) categories.

The combined ROMs found in these previous studies are in line with the current results in the Netherlands, which are presented in this study. This proves the validity of our data and the robustness of the MSRSGC classification. The observed variance in ROM between the independent studies can probably be explained by the fact that they had different cohorts that may have varied in the distribution of dignity and tumor types. Metastatic cutaneous squamous cell carcinoma (cSCC) to the parotid, for instance, is one of the tumors that has varying prevalence per geographical zone.[18] The majority of the studies included in Table 6 include metastatic cSCC to the parotid. Because of the relatively easy diagnosis of metastatic cSCC on FNAC, this might positively affect the results in countries where cSCC is more often diagnosed. Because we calculated the combined ROMs, we have adjusted for this possible bias.

The combined ROMs of the previous studies and the observed ROMs in our study are slightly different from the ROMs presented by Faquin et al.[4] The ROMs of the nondiagnostic category in both this study and the combined previous studies (12.5% and 14.1%) proved to be lower than the rate of 25% proposed by the authors of the MSRSGC. For the AUS category, both our results and the combined results from previous studies (29% and 34.8%) showed higher ROMs than proposed by the original authors (20%). Lastly, the ROMs of the suspected malignant category (current study, 83%; combined

studies, 80.0%) and the malignant category (current study, 99.3%; combined studies, 97.5%) were seemingly higher than the values of 60% and 90% proposed in the original classification. Therefore, we propose changing the expected ROMs of the categories as follows: <15% for the nondiagnostic category, ±30% for the AUS category, >80% for the suspected malignant category, and >95% for the malignant category. These changes, in our opinion, do not affect the original management propositions of the MSRSGC classification.

The sensitivity of FNAC was lower in GHs and HNOACs than HNOCs, whereas the specificity was nearly the same. This implicates a higher chance of false-negative results (and thus missing the diagnosis of malignancy) at less specialized hospitals. There are 2 likely causes for these differences. First, it is well established that the procedure's accuracy is lower when the operators who perform the fine-needle aspiration (H&N surgeons, radiologists, or pathologists) are less experienced.[19] Second, the combination of a relatively rare diagnosis such as salivary gland carcinoma and the previously mentioned pitfalls in cytopathology (different tumor types with overlapping in morphology and cell types with only subtle or no differences between tumors, limited benefit from additional staining, and a lack of clear cytomorphological signs of malignancy in some malignant tumors) make the pathologist's experience in the assessment of salivary gland FNAC material also likely to influence the diagnostic accuracy. However, we have to note that even the most experienced H&N pathologists can find the diagnosis of some salivary gland tumors, on both cytopathology and histopathology, troublesome.

The histopathological subtypes that were most often misdiagnosed on FNAC were myoepithelial carcinoma, epithelial-myoepithelial carcinoma, mucoepidermoid carcinoma, carcinoma ex pleomorphic adenoma, acinic cell carcinoma, and adenoid cystic carcinoma. Myoepithelial carcinoma, adenoid cystic carcinoma, and epithelial-myoepithelial carcinoma show a resemblance to pleomorphic adenomas because of shared cell types, and they often lack overt features of malignancy. Acinic cell carcinoma is often mistaken on cytopathology for normal salivary gland tissue because it can show a close resemblance to normal acinic cells. The difficulty in the correct cytopathological diagnosis of (low-grade) mucoepidermoid carcinoma is mostly due to sampling error because this tumor type consists of both an epithelial component and a cystic component. If only or mainly the cystic component is sampled, the tumor can be mistaken for a mucous cyst or Warthin tumor, whereas the epithelial component contains a variety of histopathological patterns, which may show overlapping with pleomorphic adenoma.[20,21] Carcinoma ex pleomorphic adenoma is also prone to false-negative results, partly because of an overlap in morphological components with other tumor types and partly because of a considerable risk of sampling error when only the benign segment of the tumor is sampled.[10,22]

Also, there were several cytopathological diagnoses carrying a relatively high risk for a false-negative result. Although the total number of diagnoses was small, 50% of the preoperatively assumed basal cell adenomas and 33.3% of the assumed myoepitheliomas turned out to be malignant. This can be explained by the fact that invasive growth, which is the main discriminator between these two and their malignant counterparts (basal cell carcinoma and myoepithelial carcinoma), is never clear on cytopathology. Furthermore, oncocytoma (9.1%) showed relatively high false-negative rates. Clinicians should, therefore, be warned in case of these 3 cytopathological diagnoses. Cytopathologically diagnosed Warthin tumors and pleomorphic adenomas both carry a low false-negative rate of ±2%. The majority of false-negatively diagnosed Warthin tumors were primary salivary gland malignancies such as mucoepidermoid carcinoma and acinic cell carcinoma. The first was probably caused by the cystic components of Warthin tumors, which are also found in low-grade mucoepidermoid carcinoma, along with the general resemblance of their cell types. The reason that acinic cell carcinoma is often mistaken for a Warthin tumor lies mainly in the prominent (cystic) lymphoid infiltrate seen in both.

High-volume surgery can have a favorable effect on overall survival in the treatment of H&N cancers.[23-25] However, high volumes are not easily achieved in the treatment of rare cancers such as salivary gland carcinoma, let alone for all its distinct subtypes. Centralization of care, therefore, is of the utmost importance for achieving high volumes. Previous studies have shown that the centralization of care for rare diseases can be beneficial to overall survival.[26] Likewise, a recent study showed that major salivary gland carcinoma could also benefit from centralization of care because high-volume hospitals were shown to have lower rates of positive surgical margins.[27] Therefore, it is our strong belief that malignant parotid
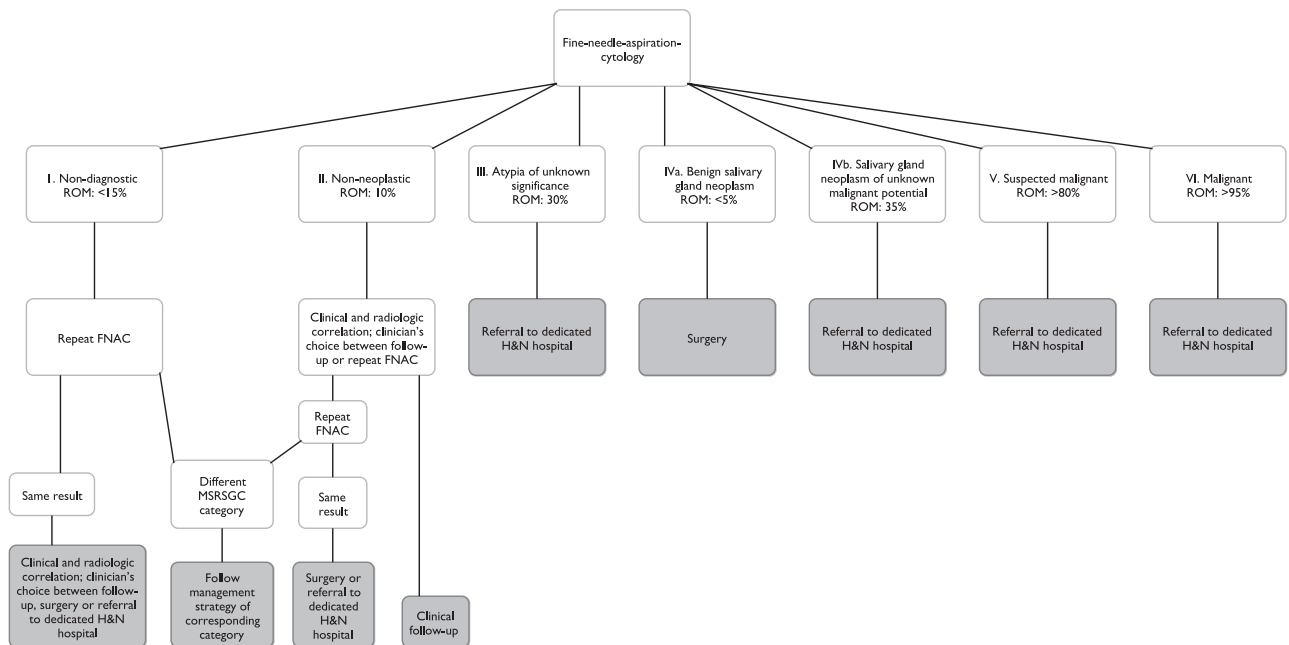
Figure 2. Management and referral flowchart for general hospitals for cases of suspected parotid gland neoplasms. FNAC indicates fine-needle aspiration cytology; H&N, head and neck; MSRSGC, Milan System for Reporting Salivary Gland Cytopathology; ROM, risk of malignancy.

tumors should be surgically removed at dedicated high-volume centers. Unfortunately, our results showed that many resections of malignant parotid gland tumors were performed in GHs (15.4%).

In most of the malignant parotid tumors resected at a GH, the patient was not referred to a dedicated H&N center because of a false-negative preoperative FNAC (MSRSGC I, II, or IVa was found in 53%). However, 17.8% had a preoperative diagnosis of SUMP (MSRSGC IVb), and 4.9% had a preoperative diagnosis of AUS (MSRSGC III). In 24.3%, the preoperative diagnosis even was malignant or suspicious for malignancy (MSRSGC V or VI).

Because of the high numbers of surgically removed malignant tumors at GHs, the previously presented ROMs, and the differences in the false-negative rates and sensitivity among the various treatment facilities, a referral scheme for clinicians at GHs is proposed. The goal is to minimize the chance of false-negative preoperative cytopathological diagnoses and to ensure that more patients with malignant parotid tumors are treated at a dedicated H&N center.

In summary, a nondiagnostic result (MSRSGC I) should warrant repeating FNAC. In cases with the same (nondiagnostic) result, clinical and radiological correlation is necessary. When there is no or low suspicion of a neoplasm, follow-up can be appropriate. If there is a high suspicion of neoplastic disease, surgery (when a benign tumor is suspected) or referral to a dedicated H&N center (in case of clinical or radiological doubt about the dignity) should be considered. Nonneoplastic results (MSRSGC II) should be carefully evaluated. The clinician must decide after clinical and radiological correlation if follow-up or repeat FNAC is indicated. If this repeated FNAC yields similar results and the FNAC result does not correspond to the clinical and/or radiological correlation, surgery (when a benign tumor is suspected) or referral to a dedicated H&N center is indicated. In patients with a cytopathological diagnosis of AUS (MSRSGC III), referral to a dedicated H&N hospital should be considered on the basis of the ROM of 30%. Clinicians at dedicated centers are encouraged to repeat the FNAC because of the higher diagnostic accuracy of FNAC at these centers. Benign salivary gland neoplasms (MSRSGC IVa) can be removed safely by an experienced surgeon at a GH. Because of the relatively high ROM (35%) of SUMP (MSRSGC IVb), treating these neoplasms at a dedicated H&N center should be considered. Clinicians at dedicated centers are encouraged to repeat the FNAC because of the higher diagnostic accuracy of FNAC at these centers. We strongly

advise that suspected malignant and malignant tumors (MSRSGC V and VI) be treated at a dedicated H&N hospital. This referral scheme is summarized in the flowchart in Figure 2.

On account of the sole inclusion of patients who had a histopathological correlation after their parotid gland FNAC, this study may be limited by the underrepresentation of FNACs in MSRSGC II (nonneoplastic) because these are not always surgically resected. Furthermore, earlier studies have observed that lymphomas predominantly contribute to a higher ROM in the nonneoplastic category.[5] These were also excluded from our analysis because they were underrepresented in this surgically treated cohort. Another limiting factor is that no histopathological reassessment of the resections was performed. A previous study has shown that the reassessment of histopathology in malignant salivary gland tumors is associated with changes in the histopathological subtype, the origin of the tumor, or even dignity in a number of patients.[28] Because of our large sample size, histopathological reassessment was practically impossible to perform.

Our study, which to our knowledge is the most extensive retrospective, nationwide study evaluating parotid gland FNACs to date, once more proves the effectiveness and validity of the MSRSGC. Therefore, we strongly recommend the use of this diagnostic tool for reporting on salivary gland cytology. In addition, we present referral guidelines for clinicians at GHs based on our results. These help in minimizing the chance of false-negative preoperative FNAC results and can further optimize care for patients with parotid gland tumors.

FUNDING SUPPORT

CONFLICT OF INTEREST DISCLOSURES

AUTHOR CONTRIBUTIONS
Sam T. H. Reerds: Study concept, methodology design, data collection, statistical analysis, supervision of the statistical analysis, writing–original draft, drafting of the management and referral flowchart for general hospitals, and revision and finalization of the manuscript. **Adriana C. H. Van Engen–Van Grunsven:** Study concept, methodology design, supervision of the statistical analysis, drafting of the management and referral flowchart for general hospitals, and revision and finalization of the manuscript. **Frank J. A. van den Hoogen:** Drafting of the management and referral flowchart for general hospitals and revision and finalization of the manuscript. **Robert P. Takes:** Drafting of the management and referral flowchart for general hospitals and revision and finalization of the manuscript. **Henri A. M. Marres:** Study concept, methodology design, drafting of the management and referral flowchart for general hospitals, and revision and finalization of the manuscript. **Jimmie Honings:** Study concept, methodology design, supervision of the statistical analysis, writing–original draft, drafting of the management and referral flowchart for general hospitals, and revision and finalization of the manuscript.

REFERENCES

1. El-Naggar AK, Chan JKC, Grandis JR, Takata T, Slootweg PJ, eds. WHO Classification of Head and Neck Tumours. 4th ed. IARC; 2017.
2. Liu CC, Jethwa AR, Khariwala SS, Johnson J, Shin JJ. Sensitivity, specificity, and posttest probability of parotid fine-needle aspiration: a systematic review and meta-analysis. *Otolaryngol Head Neck Surg*. 2016;154:9-23. doi:10.1177/0194599815607841
3. Rossi ED, Faquin WC, Baloch Z, et al. The Milan System for Reporting Salivary Gland Cytopathology: analysis and suggestions of initial survey. *Cancer Cytopathol*. 2017;125:757-766. doi:10.1002/cncy.21898
4. Faquin WC, Rossi ED, Baloch Z, et al, eds. The Milan System for Reporting Salivary Gland Cytopathology. Springer; 2018:chap 1-9.
5. Viswanathan K, Sung S, Scognamiglio T, Yang GCH, Siddiqui MT, Rao RA. The role of the Milan System for Reporting Salivary Gland Cytopathology: a 5-year institutional experience. *Cancer Cytopathol*. 2018;126:541-551. doi:10.1002/cncy.22016
6. Savant D, Jin C, Chau K, et al. Risk stratification of salivary gland cytology utilizing the Milan system of classification. *Diagn Cytopathol*. 2019;47:172-180. doi:10.1002/dc.24063
7. Dubucs C, Basset C, D'Aure D, Courtade-Saidi M, Evrard SM. A 4-year retrospective analysis of salivary gland cytopathology using the Milan System for Reporting Salivary Gland Cytology and ancillary studies. *Cancers (Basel)*. 2019;11:1912. doi:10.3390/cancers11121912
8. Mazzola F, Gupta R, Luk PP, Palme C, Clark JR, Low TH. The Milan System for Reporting Salivary Gland Cytopathology—proposed modifications to improve clinical utility. *Head Neck*. 2019;41:2566-2573. doi:10.1002/hed.25732
9. Wu HH, Alruwaii F, Zeng BR, Cramer HM, Lai CR, Hang JF. Application of the Milan System for Reporting Salivary Gland Cytopathology: a retrospective 12-year bi-institutional study. *Am J Clin Pathol*. 2019;151:613-621. doi:10.1093/ajcp/aqz006
10. Song SJ, Shafique K, Wong LQ, LiVolsi VA, Montone KT, Baloch Z. The utility of the Milan system as a risk stratification tool for salivary gland fine needle aspiration cytology specimens. *Cytopathology*. 2019;30:91-98. doi:10.1111/cyt.12642
11. Choy KCC, Bundele MM, Li H, Fu EW, Rao NCL, Lim MY. Risk stratification of fine-needle aspiration cytology of parotid neoplasms based on the Milan system—experience from a tertiary center in Asia. *Head Neck*. 2019;41:3125-3132. doi:10.1002/hed.25804
12. Chen YA, Wu CY, Yang CS. Application of the Milan System for Reporting Salivary Gland Cytopathology: a retrospective study in a tertiary institute. *Diagn Cytopathol*. 2019;47:1160-1167. doi:10.1002/dc.24279
13. Lee JJL, Tan HM, Chua DYS, Chung JGK, Nga ME. The Milan system for reporting salivary gland cytology: a retrospective analysis of 1384 cases in a tertiary Southeast Asian institution. *Cancer Cytopathol*. 2020;128:348-358. doi:10.1002/cncy.22245
14. Mazzola F, Tomasoni M, Mocellin D, et al. A multicenter validation of the revised version of the Milan System for Reporting Salivary Gland Cytology (MSRSGC). *Oral Oncol*. 2020;109:104867. doi:10.1016/j.oraloncology.2020.104867

15. Pujani M, Chauhan V, Agarwal C, Raychaudhuri S, Singh K. A critical appraisal of the Milan System for Reporting Salivary Gland Cytology (MSRSGC) with histological correlation over a 3-year period: Indian scenario. *Diagn Cytopathol.* 2019;47:382-388. doi:10.1002/dc.24109

16. Casparie M, Tiebosch AT, Burger G, et al. Pathology databanking and biobanking in the Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. *Cell Oncol.* 2007;29:19-24. doi:10.1155/2007/971816

17. Eveson J, Auclair PL, Gnepp DR, El Naggar AK. Tumours of the salivary gland. In: Barnes L, Eveson JW, Reichart P, Sidransky D, eds. World Health Organization Classification of Tumours: Pathology and Genetics of Head and Neck Tumours. IARC; 2005:164.

18. Leiter U, Garbe C. Epidemiology of melanoma and nonmelanoma skin cancer—the role of sunlight. *Adv Exp Med Biol.* 2008;624:89-103. doi:10.1007/978-0-387-77574-6_8

19. Jandu M, Webster K. The role of operator experience in fine needle aspiration cytology of head and neck masses. *Int J Oral Maxillofac Surg.* 1999;28:441-444.

20. Joseph TP, Joseph CP, Jayalakshmy PS, Poothiode U. Diagnostic challenges in cytology of mucoepidermoid carcinoma: report of 6 cases with histopathological correlation. *J Cytol.* 2015;32:21-24. doi:10.4103/0970-9371.155226

21. Vasudevan G, Bishnu A, Singh BMK, Singh VK. Mucoepidermoid carcinoma of salivary gland: limitations and pitfalls on FNA. *J Clin Diagn Res.* 2017;11:ER04-ER06. doi:10.7860/JCDR/2017/25341.9941

22. Antony J, Gopalan V, Smith RA, Lam AK. Carcinoma ex pleomorphic adenoma: a comprehensive review of clinical, pathological and molecular data. *Head Neck Pathol.* 2012;6:1-9. doi:10.1007/s12105-011-0281-z

23. Eskander A, Irish J, Groome PA, et al. Volume-outcome relationships for head and neck cancer surgery in a universal health care system. *Laryngoscope.* 2014;124:2081-2088. doi:10.1002/lary.24704

24. Wuthrick EJ, Zhang Q, Machtay M, et al. Institutional clinical trial accrual volume and survival of patients with head and neck cancer. *J Clin Oncol.* 2015;33:156-164. doi:10.1200/JCO.2014.56.5218

25. de Ridder M, Balm AJM, Baatenburg de Jong RJ, et al. Variation in head and neck cancer care in the Netherlands: a retrospective cohort evaluation of incidence, treatment and outcome. *Eur J Surg Oncol.* 2017;43:1494-1502. doi:10.1016/j.ejso.2017.02.017

26. Weitz J, Koch M, Friess H, Buchler MW. Impact of volume and specialization for cancer surgery. *Dig Surg.* 2004;21:253-261. doi:10.1159/000080198

27. Bollig CA, Zitsch RP III. Impact of treating facilities' type and volume in patients with major salivary gland cancer. *Laryngoscope.* 2019;129:2321-2327. doi:10.1002/lary.27844

28. Stodulski D, Majewska H, Skalova A, Mikaszewski B, Biernat W, Stankiewicz C. Histological reclassification of parotid gland carcinomas: importance for clinicians. *Eur Arch Otorhinolaryngol.* 2016;273:3937-3942. doi:10.1007/s00405-016-4048-8