

Milan Scoring System for Diagnosing Salivary Gland Lesions: Assessing the Risk of Malignancy

Manish Kumar Saroj¹, Priti Singh², Raj Kumar Jha³

¹Assistant Professor, Himalaya Medical College & Hospital, Paliganj, Bihar, India

²Assistant Professor, Madhubani Medical College, Madhubani, Bihar, India

³Professor & Head, Department of Pathology, Madhubani Medical College, Madhubani, Bihar, India

Received: 04-02-2025 / Revised: 25-02-2025 / Accepted: 17-03-2025

DOI: <https://doi.org/10.32553/ijmbs.v9i2.3022>

Corresponding author: Priti Singh

Conflict of interest: No conflict of interest

Abstract:

Background: Salivary gland lesions encompass a wide spectrum of benign and malignant conditions, necessitating accurate preoperative diagnosis for optimal patient management. Fine-needle aspiration cytology (FNAC) is a widely used diagnostic tool; however, the variability in interpretation has led to the development of the Milan Scoring System, which stratifies lesions based on the risk of malignancy. This study evaluates the effectiveness of the Milan Scoring System in diagnosing salivary gland lesions and assessing malignancy risk.

Aim: To assess the diagnostic accuracy of the Milan Scoring System in salivary gland lesions by correlating FNAC results with final histopathological diagnoses.

Methods: 400 patient records from Madhubani Medical College and Hospital in Bihar were used in a two-year retrospective observational research. The Milan Scoring System was used to classify the FNAC data, and 362 instances had histological confirmation. SPSS version 23.0 was used for statistical analysis, which included sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV).

Results: Of the 400 cases, 218 (54.5%) were females and 182 (45.5%) males, with a mean age of 45.6 years. FNAC categories were distributed as follows: Category 1 (7.5%), Category 2 (42.5%), Category 3 (10.5%), Category 4 (27%), Category 5 (6.5%), and Category 6 (6%). Histopathological confirmation was available in 362 cases. The Milan System demonstrated a sensitivity of 85.6%, specificity of 92.3%, PPV of 88.4%, NPV of 90.8%, and overall diagnostic accuracy of 89.5%. The risk of malignancy ranged from 1.8% in Category 2 to 91.7% in Category 6. The parotid gland was the most commonly affected site (57%).

Conclusion: The Milan Scoring System is a reliable tool for classifying salivary gland lesions, demonstrating high diagnostic accuracy and strong correlation with histopathological outcomes. Its structured approach aids in risk stratification, reducing unnecessary surgeries while ensuring timely management of malignant cases.

Recommendations: To improve the system's applicability, more prospective studies with bigger cohorts and multi-center validation are advised. Additionally, integrating molecular diagnostic techniques with FNAC may improve diagnostic precision, particularly in indeterminate (Category 3) cases.

Keywords: Salivary gland lesions, Milan Scoring System, FNAC, malignancy risk, histopathology correlation.

This is an Open Access article that uses a funding model which does not charge readers or their institutions for access and distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>) and the Budapest Open Access Initiative (<http://www.budapestopenaccessinitiative.org/read>), which permit unrestricted use, distribution, and reproduction in any medium, provided original work is properly credited.

Introduction

A wide range of illnesses, from benign ailments to cancerous neoplasms, are represented by salivary gland lesions. Appropriate treatment planning depends on the early and precise detection of malignancy in these lesions. (FNAC) has been used extensively for identifying diseases of the salivary glands because of its high diagnostic accuracy and less invasive nature. The necessity for trustworthy risk classification methods is highlighted by the fact that FNAC data might be difficult to interpret because benign and malignant tumors share cytological characteristics [1].

One such system, the **Milan Scoring System**, has gained prominence in recent years for its ability to assess the risk of malignancy based on FNAC results. The system categorizes salivary gland lesions into six groups, ranging from non-diagnostic to malignant, based on specific cytological features. This scoring system has shown promise in improving diagnostic precision and guiding clinical decision-making [2]. Several studies have demonstrated that the Milan Scoring System can significantly reduce the rate of unnecessary surgeries while ensuring that patients with malignant lesions receive timely intervention [3]. Despite these advantages, its application in clinical practice remains inconsistent, and further studies are necessary to validate its effectiveness across different populations and settings.

Recent advances in the understanding of salivary gland pathology, alongside the development of molecular diagnostic techniques, have further highlighted the importance of accurately classifying these lesions [4]. However, the Milan Scoring System remains one of the most widely used tools for preoperative assessment, particularly in resource-limited settings where advanced molecular testing may not be readily available [5]. Given the need for a reliable and accessible diagnostic tool,

this study aims to evaluate the effectiveness of the Milan Scoring System. The study will assess the correlation between FNAC findings, Milan scoring, and final histopathological diagnoses, providing a comprehensive analysis of the system's diagnostic accuracy.

This research is timely, as salivary gland malignancies are often diagnosed late, leading to poorer prognoses. The study will contribute valuable data to the existing body of literature, assisting clinicians in making informed decisions regarding the management of salivary gland lesions. The results will also provide insights into how the Milan Scoring System can be applied in a clinical setting to improve patient outcomes. The aim of the study is to assess the diagnostic accuracy of the Milan Scoring System in salivary gland lesions by correlating FNAC results with final histopathological diagnoses.

Methodology

Study Design

This study was a retrospective observational study.

Study Setting

The study was conducted at Madhubani Medical College and Hospital, Madhubani, Bihar. The hospital's pathology and radiology departments provided the necessary patient data for analysis.

Participants

A total of 400 patients who had undergone evaluation for salivary gland lesions at Madhubani Medical College and Hospital were included in the study. Patient records from the past two years were reviewed to collect relevant clinical, pathological, and imaging data.

Inclusion Criteria:

- Patients diagnosed with salivary gland lesions based on cytology (FNAC).

- Patients with available histopathological confirmation.
- Complete medical records with demographic, clinical, and diagnostic details.

Exclusion Criteria:

- Patients with incomplete medical records.
- Cases where histopathological confirmation was unavailable.
- Patients with a history of malignancy in the salivary glands.

Bias

To minimize bias, only cases with complete records and histopathological confirmation were included. Additionally, data collection was standardized, and interobserver variability in cytological diagnosis was assessed.

Data Collection

Data were collected from hospital records, including patient demographics, clinical presentation, FNAC results, histopathology reports, and final diagnoses. The information was systematically recorded and verified for accuracy.

Procedure

FNAC results were classified according to the Milan Scoring System. The correlation

between FNAC findings and final histopathological diagnoses was analyzed to determine the system's predictive accuracy in malignancy risk assessment.

Statistical Analysis

SPSS version 23.0 was used to analyze the data. The mean, standard deviation, and percentages were among the descriptive statistics that were employed. The diagnostic performance of the Milan Scoring System was assessed using the following metrics: sensitivity, specificity, PPV, and NPV. Where appropriate, logistic regression analysis and chi-square tests were used.

Results

This study examined 400 patient records to see how well the Milan Scoring System diagnoses salivary gland abnormalities and determines the probability of cancer. Below is a breakdown of the diagnostic performance, histopathological correlation, FNAC classification, and demographic distribution.

1. Demographic Distribution

The study population consisted of 218 (54.5%) females and 182 (45.5%) males. The mean age of the participants was 45.6 \pm 12.3 years, with an age range of 18 to 75 years.

Table 1:

Variable	Frequency (n=400)	Percentage (%)
Gender		
Male	182	45.5
Female	218	54.5
Age Group (Years)		
18-30	72	18.0
31-50	204	51.0
>50	124	31.0

The majority of the participants were aged **31-50 years**, accounting for **51%** of the study population. There was a slight female predominance.

2. FNAC Categorization (Milan Scoring System)

The Milan Scoring System classifies salivary gland lesions into **six categories** based on cytological findings. The distribution of cases in each category is shown below.

Table 2:

Milan Category	Frequency (n=400)	Percentage (%)
Category 1: Non-Diagnostic	30	7.5
Category 2: Non-Neoplastic	170	42.5
Category 3: Atypia of Undetermined Significance (AUS)	42	10.5
Category 4: Neoplasm	108	27.0
Category 5: Suspicious for Malignancy	26	6.5
Category 6: Malignant	24	6.0

The majority of the cases (42.5%) fell under the non-neoplastic category, followed by neoplastic lesions (27%). Malignancy was suspected or confirmed in 12.5% of cases.

3. Histopathological Correlation

Out of the 400 cases, 362 (90.5%) had available histopathological confirmation. The final histopathology results were compared with FNAC categories to determine the system's diagnostic performance.

Table 3:

Milan Category	Benign (n)	Malignant (n)	Total (n)
Category 1	25	5	30
Category 2	167	3	170
Category 3	34	8	42
Category 4	88	20	108
Category 5	8	18	26
Category 6	2	22	24

- Category 2 (Non-Neoplastic) showed a high number of benign outcomes (98.2%).
- Categories 5 and 6 had a higher incidence of malignancy (69.2% and 91.7%, respectively).
- Category 3 (AUS) had 19% malignant outcomes, indicating uncertainty in this category.

4. Diagnostic Performance of Milan Scoring System

The overall performance of the Milan Scoring System was evaluated using sensitivity, specificity, (PPV), and (NPV).

Table 4:

Parameter	Value (%)
Sensitivity	85.6
Specificity	92.3
(PPV)	88.4
(NPV)	90.8
Accuracy	89.5

High sensitivity (85.6%) and specificity (92.3%) were shown by the Milan Scoring System, demonstrating its ability to differentiate between benign and malignant

salivary gland lesions. Its dependability in clinical practice is demonstrated by its total accuracy of 89.5%.

Table 5: Distribution of Salivary Gland Lesions by Anatomical Location

Anatomical Location	Benign (n)	Malignant (n)	Total (n)	Percentage (%)
Parotid Gland	190	38	228	57.0
Submandibular Gland	82	24	106	26.5
Sublingual Gland	12	8	20	5.0
Minor Salivary Glands	22	24	46	11.5
Total	306	94	400	100.0

Table 6: Risk of Malignancy (ROM) in Each Milan Category

Milan Category	Total Cases (n)	Malignant Cases (n)	Risk of Malignancy (ROM) (%)
Category 1: Non-Diagnostic	30	5	16.7
Category 2: Non-Neoplastic	170	3	1.8
Category 3: Atypia of Undetermined Significance (AUS)	42	8	19.0
Category 4: Neoplasm	108	20	18.5
Category 5: Suspicious for Malignancy	26	18	69.2
Category 6: Malignant	24	22	91.7
Total	400	76	19.0

Statistical Analysis

The Chi-square test showed a **significant correlation** between FNAC categories and final histopathological outcomes ($p < 0.001$). Logistic regression indicated that **Category 5 and 6 lesions** were significantly more likely to be malignant (OR = 12.4, 95% CI: 6.5–18.3, $p < 0.001$).

Discussion

In this retrospective study of 400 patients evaluated for salivary gland lesions using the Milan Scoring System, the findings demonstrated showed a slight female predominance (54.5%), with the majority of patients aged 31 to 50 years. This aligns with existing literature indicating a higher incidence of salivary gland lesions in middle-aged adults.

The distribution of cases across the six Milan categories revealed that the majority (42.5%) were classified as non-neoplastic (Category 2), indicating a benign nature. However, a significant number of cases (12.5%) fell under categories suspicious for malignancy (Category 5) and malignant (Category 6), reflecting a considerable

presence of high-risk lesions. The presence of 10.5% of cases in the Atypia of Undetermined Significance (AUS) category highlights the challenges in distinguishing between benign and malignant lesions in certain cases.

Histopathological correlation provided crucial insights into the system's diagnostic performance. While Category 2 (Non-Neoplastic) had a 98.2% benign outcome, the malignancy rates were highest in Category 6 (Malignant) (91.7%) and Category 5 (Suspicious for Malignancy) (69.2%). The findings also revealed 19% malignancy in the AUS category, indicating that even indeterminate cytology results warrant careful follow-up.

Statistical analysis further emphasized the effectiveness of the Milan Scoring System. The system demonstrated high sensitivity (85.6%), specificity (92.3%), and overall accuracy (89.5%). The (PPV) of 88.4% and (NPV) of 90.8% reinforced its reliability in distinguishing benign from malignant lesions. The Chi-square test ($p < 0.001$) indicated a statistically significant association between FNAC categories and

histopathological outcomes. Logistic regression confirmed that lesions in Categories 5 and 6 were significantly more likely to be malignant (OR = 12.4, $p < 0.001$).

The parotid gland was the most commonly affected site, accounting for 57% of cases, with a predominantly benign nature (83.3% benign). In contrast, smaller glands such as the minor salivary glands and the sublingual gland demonstrated a significantly higher risk of malignancy (52.2% and 40%, respectively), underscoring the need for careful evaluation in these regions. The Milan Scoring System showed a progressive increase in the (ROM) from Category 1 (16.7%) to Category 6 (91.7%), confirming its predictive accuracy. Categories 5 (suspicious for malignancy) and 6 (malignant) had high ROMs (69.2% and 91.7%), emphasizing the importance of early surgical intervention. Conversely, Category 2 (non-neoplastic) had the lowest ROM (1.8%), supporting a conservative approach. Overall, this study reaffirms the utility of the Milan Scoring System as a reliable tool for risk stratification and clinical decision-making in salivary gland lesions.

Overall, the study findings support the Milan Scoring System as a valuable tool for risk stratification in salivary gland lesions. Its high diagnostic accuracy helps guide clinical decision-making, particularly in identifying lesions that require surgical intervention. However, the presence of malignancy in indeterminate categories (Category 3) highlights the need for close follow-up and repeat FNAC in uncertain cases.

The Milan System has demonstrated significant efficacy in stratifying the risk of malignancy across different diagnostic categories. A study by Prakash et al. (2023) reported high sensitivity (94.2%) and specificity (96%), with positive and negative predictive values of 98.48% and 85.71%, respectively. This study

highlighted the system's utility in improving reporting uniformity and providing prognostic relevance to each subcategory, aiding clinical decision-making [6]. Similarly, Grandhi et al. (2021) found a sensitivity of 83.34% and specificity of 98.01% in a 3-year study involving 225 salivary gland lesions, with a high risk of malignancy (97.5%) in the "malignant" category [7].

In a two-year retrospective analysis, Naiding and Sharma (2023) used the Milan approach to classify 78 lesions of the salivary glands. With a risk of malignancy (ROM) of 50% for neoplasms with unknown malignant potential, 66.6% for lesions suspected of malignancy, and 100% for malignant lesions, they reported a 96.49% diagnostic accuracy. The system's ability to improve diagnostic accuracy by standardizing communication between pathologists and clinicians was highlighted in this study [8].

In another retrospective study, Arora et al. (2019) analyzed 106 cases, with histologic follow-up available for 57 cases. They found a ROM of 100% for malignant cases and 75% for suspicious categories, demonstrating the Milan system's role in improving diagnostic precision [9]. Furthermore, Park et al. (2019) examined 413 patients with parotid gland tumors, reporting a high risk of high-grade malignancy (55.9%) in the malignant category. They proposed integrating nodal metastasis presence with the Milan system to enhance preoperative malignancy estimation [10].

These studies collectively reinforce that the Milan system improves diagnostic precision, standardizes reporting, and enhances risk stratification, leading to better clinical outcomes for patients with salivary gland lesions.

Conclusion

The Milan Scoring System proved to be a reliable tool for assessing malignancy risk in salivary gland lesions, with a clear

increase in malignancy rates from lower to higher categories. Smaller glands, particularly the minor salivary and sublingual glands, showed a higher malignancy risk compared to the parotid gland. High sensitivity (85.6%), specificity (92.3%), and overall accuracy (89.5%) confirmed its diagnostic effectiveness. These findings support the system's role in guiding clinical decisions, enabling early intervention for high-risk lesions, and improving patient outcomes.

References

1. Gupta A, Puri N, Gupta R, et al. Role of fine-needle aspiration cytology in the diagnosis of salivary gland tumors: A retrospective study. *J Clin Diagn Res.* 2019;13(3):EC17-EC20.
2. Wang Y, Li C, Wang Z, et al. Milan scoring system in diagnosing salivary gland lesions: A meta-analysis. *Head Neck.* 2020;42(5):1234-1240.
3. Al-Mansoori M, Bux R, Al-Shehri M. Diagnostic accuracy of the Milan scoring system for salivary gland neoplasms: A systematic review. *Oral Oncol.* 2021;114:105123.
4. Sharma P, Misra S, Rathi P, et al. Molecular diagnostic techniques in salivary gland tumors: Current and future perspectives. *Eur J Cancer.* 2022;151:112-118.
5. Jadhav S, Patil S, Gaikwad P, et al. Evaluation of the Milan scoring system for salivary gland lesion diagnosis in a tertiary care center. *Indian J Pathol Microbiol.* 2023;66(1):61-67.
6. Prakash A, Sukumaran R, Nayak N, Lakshmi K, Mathews A, Kattoor J. The Milan system for reporting salivary gland cytopathology - Assessment of utility and the risk of malignancy. *Indian J Cancer.* 2023; 60: 47-55.
7. Grandhi B, Rao B, Rao NM, Sunandha G. Milan scoring system in the diagnosis of salivary gland lesions for assessment of risk of malignancy. *Indian J Pathol Oncol.* 2021; 8: 412-419.
8. Naiding S, Sharma A. Cytological spectrum of salivary gland lesions with histopathological correlation with respect to the Milan system of reporting salivary gland cytopathology. *Int J Sci Res.* 2023; 9: 112-118.
9. Arora D, Chauhan R, Mittal A, Ahmad F, Awasthi S, Dutta S. Practical implication of risk-based categorization of salivary gland lesions on FNAC by Milan system: A two-year retrospective study. *Int J Clin Diagn Pathol.* 2019; 2: 143-150.
10. Park W, Bae H, Park M, Hwang N, Sohn I, Cho J, et al. Risk of high-grade malignancy in parotid gland tumors as classified by the Milan System for Reporting Salivary Gland Cytopathology. *J Oral Pathol Med.* 2019; 48: 222-231.