Correlation between Yokohama Cytological Coding and Radiological Findings and Their Diagnostic Accuracies against Histopathology: A Retrospective Study of Palpable Breast Lesions

Puja Singh1*, Bharti Badlani2, Chanchalesh Dehariya2, Munira Murtaza Joher3

1Bundelkhand Government Medical College, Sagar, India
2Chhindwara Institute of Medical Sciences, Chhindwara, India
3Bradford Teaching Hospital, Bradford, United Kingdom

Background: Breast carcinoma is the most common malignancy and demands quick and accurate diagnosis and treatment. Precise diagnosis of breast lesions is made using a triple-test approach: clinical, radiological and cytological. However, multiple steps make the process time-consuming and expensive. In developing countries like India, trained and certified radiologists are extremely overburdened. Fine needle aspiration cytology (FNAC) along with clinical examination can fill the gap. This study aims to correlate cytological, radiological and histological findings and measure their relative accuracies. Based on these findings, a new approach will be proposed to address the above shortcomings.

Materials and methods: The FNAC was performed on all cases and reported as per Yokohama cytology. The cytological findings were correlated & validated against radiological and histopathological findings respectively. Relative performance of cytological and radiological findings were established using sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy. A chi-Square test for independence between cytological and radiological findings was performed.

Results: The sensitivity, specificity, PPV, NPV, and accuracy for cytological findings come out as 97.60, 90.08, 90.37, 97.52, and 93.75, respectively. Meanwhile, the radiological findings come out as 96.61, 82.20, 84.44, 96.04, and 89.41, respectively. The chi-square test demonstrates strong interdependence between cytological and radiological findings.

Conclusion: FNAC is more accurate, quicker, and cheaper than Radiological tests. Hence, FNAC based on the Yokohama system, along with clinical observations, can be used as a primary diagnosis tool in developing countries with limited health resources without making significant compromises on incorrect treatment. If needed, radiology and histopathology can be used for precise diagnosis and treatment.

Keywords: FNAC, cytology, breast lesions, Yokohama, radiology, histopathology

Date of submission: February 2, 2024
Last Revised: May 27, 2024
Accepted for publication: June 3, 2024

Corresponding Author:
Puja Singh
Department of Pathology
Bundelkhand Government Medical College
Madhya Pradesh, 470001, Sagar, India
e-mail: dr.pujasingh@gmail.com

Copyright © 2024 Cell and BioPharmaceutical Institute. This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC-BY-NC) License.
**Introduction**

Breast cancer (BC) is the most common malignancy among women globally.\(^1,2\) It has now surpassed lung cancer as the leading cause of global cancer incidence in 2020, with an estimated 2.3 million new cases, representing 11.7% of all cancer cases. Breast carcinoma in Indian women is now the most prevalent cancer, with about 26.6% of all cancer cases, surpassing cervical cancer, with about 17.7% of all cancer cases.\(^3\) This is especially true for many urban cities.\(^4\) There is a positive correlation between the incidence of reported instances and people's awareness of breast lesions. This very fact constrains expedient diagnosis and precise therapy.\(^5,6\)

Quick diagnosis sometimes becomes very crucial for proper disease management. Traditionally, an accurate diagnosis of breast lesions is made using a triple-test approach of clinical examination, cytology and radiographic examination. As this process involves multiple steps, quick diagnosis sometimes is not possible. In addition, multiple steps make the entire procedure expensive.\(^7\)

Cytology of breast lesions is often performed via fine needle aspiration cytology (FNAC) or fine needle biopsy (FNB). The FNAC offers a quick, cheap, and minimally invasive diagnostic solution with high accuracy. However, the lack of global standards made the diagnosis inconsistent and non-repeatable. In 2016, the International Academy of Cytology put forth the Yokohama System for FNAC classification of breast lesions into 5 categories. This system not only made the diagnosis comprehensive and standardised but also clearly outlined risk of malignancy (ROM) and treatment action.\(^8\) Similarly, the American College of Radiology’s BI-RADS classified mammographic findings into 6 categories.\(^9\)

Access to medical facilities is more of a privilege in developing countries. The use of a triple-test approach for breast lesions not only increases the overall time for making the assessment but also increases the cost of treatment significantly. In India, the ratio of trained radiologists to the population is 1:100,000. This acute shortage of trained and certified radiologists in India further complicates the delivery of radio diagnostic services.\(^10,11\) Combined effect of these facts make the situation extremely precarious for certain sections of society. This is denying them basic of access to health equity.

There have been many studies performed across many institutes evaluating the accuracy of the Yokohama System since it was published. In addition, some of the studies also correlated with histopathological and cytological findings. However, very few studies have been performed which cover cytology, radiology, and histopathology correlation. This study has correlated the FNAC findings against both radiological and histopathological findings. Comparative sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy were also performed for cytological and radiological findings. The chi-square test of independence was performed between cytological and radiological findings. After establishing the strong correlation between the FNAC & radiological findings, this study aimed to propose an alternative to the triple-test approach which could be employed in resource-crunched developing countries.

**Material and methods**

**Study Design and Sample Collection**

This retrospective observational study was carried out at a tertiary care center in Madhya Pradesh, Central India. Required ethical clearance was taken from the Institutional Ethics Committee (IEC Ref. No: CIMS/Ethics Committee/2022/7073). Samples meeting the inclusion criteria were selected using purposive sampling technique from 2018 to 2022.

Inclusion criteria for subjects were patients, both men and women, of all age groups who were advised for FNAC of breast lump were included in the study. Patients with recurrent breast lumps, pregnant females with breast lumps, and lactating females with breast lumps were also included in the study. Meanwhile, patients with nipple discharge undergoing cytology and unidentified lumps were excluded.

**Cytology Analysis**

All the FNACs were performed following standard protocol using a 10 mL syringe and 23/24-gauge needle without using the plunger. The slides were stained using Giemsa and Papanicolaou stains. The cytology findings were reported as per Yokohama system for Breast lesions. Accordingly, all the cases were classified into 5 categories, these were unsatisfactory/inadequate, benign, atypia, suspicious, and Malignant.\(^6\)

**Radiology Examination**

Ultrasound examination of the breast lumps was performed applying a transducer gently. Both longitudinal and transverse scans were obtained. Four features were assessed...
during the scans: Shape (round, oval, or irregular), margin type (circumscribed or non-circumscribed), width (AP ratio greater than or equal to 1.4 or less than 1.4), and echogenicity (hyperechoic, isoechoic, or hyperechoic). Based on these four features, the radiologist diagnosed a score of BI-RAD from 0 to 6. Breast ultrasound was performed using Allengers Medical Systems MAM-Venus (Allengers, New Delhi, India).

**Histopathology Analysis**

Histopathological examination was performed on the tissues obtained post-surgery for breast lesions from core needle biopsy, wedge biopsy, lumpectomy, and modified radical mastectomy. Tissue was processed using an automated tissue analyzer and automated microtome. The ribbons were made between the thickness between 2-3 microns. Routine haematoxylin and eosin (HE) staining were conducted followed by microscopic examination. Two experienced pathologists reported slides obtained by the above process. These findings were considered as gold standard for evaluating the performance of other diagnostic methods.

**Statistical Analysis**

The cytological results were then correlated with the radiological findings and histopathological diagnosis. Comparative performance of cytological and radiological findings was established with the measurement of sensitivity, specificity, PPV, NPV, and accuracy. A chi-square test for independence was performed between cytological and radiological findings. The Statistical analysis was performed using Python (v3.9) (Python Software Foundation, Wilmington, DE, USA) and the libraries were Pandas (v1.3.5), Seaborn (v0.11.2), and SciPy(v1.7.1).

**Results**

**Demographic Distribution**

A total of 377 patients had been studied. The age group of the patients ranged were 17 to 87 years. Only 4 male patients had been reported in the study. All 4 male patients showed gynecomastia. The most number of cases were seen in the age group of 21-40 followed by 41-60 (Table 1). The least number of cases were seen in the 81-100 age group with only 8 cases, out of which, all three tests were performed for only 2 cases that were identified as malignant. For the age group of 0-20, out of 24 cases, one case was breast abscess, and for two other cases, histopathology was not performed. Fibro adenoma was the most common etiology involving this age group.

Cytological tests were performed for all 377 cases using the Yokohama reporting system. Findings for these were as follows: Inadequate 18.83% (n=71), benign 35.54% (n=134), atypia 8.49% (n=32), suspicious 17.77% (n=67), and malignant 19.36% (n=73). Radiological tests were performed only for 75.06% (n=283) cases. Fewer numbers were due to various reasons such as patients not willing, treating physicians not demanding it, etc. Findings for these were: Inadequate 1.86% (n=7), benign 28.12% (n=106), atypia 10.61% (n=40), suspicious 17.24% (n=65), and malignant 17.24% (n=65). Histopathological tests were performed only for 67.90% (n=256) cases. Fewer numbers were due to various reasons such as patient did not willing to undergo surgical procedure, treating physicians not demanding it, etc. Findings for these were as follows: Inadequate 0% (n=0), benign 34.75% (n=131), atypia 6.37% (n=24), suspicious 0% (n=0), and malignant 26.79% (n=101) (Table 2).

**Cytological and Radiological Correlation**

A violin plot can unequivocally present the frequency distribution across classifications. Thus, gross graphical correlation of the findings of cytological, radiological, and histopathological methods was depicted by Violin plot (Figure 1). This plot only considered 240 cases for which all three tests were performed. From this graph, we could confidently interpret a strong correlation between radiological and cytological findings, especially for malignant and benign cases with high accuracy. It could also be interpreted that there exists a strong correlation for Atypia cases.
Heat maps were used to establish the correlations of findings pairing two tests at a time (Figure 2). It could be easily inferred that a strong correlation exists between cytological and radiological findings. Out of 283 cases, a concordance of 86.92% (n=246) was found. All kinds of cases showed a good level of congruence. Out of 256 cases, the concordance between cytological and histopathological findings were 79.29% (n=203). Benign, atypia, and malignant cases showed a good level of congruence. Out of 236 cases, concordance between radiological and histopathological findings were 70.33% (n=166). Benign and malignant cases showed a good level of congruence. The correlation between cytological and radiological findings was highest among these pairs followed by cytological and histopathological findings.

**Diagnostic Accuracies**

The performance of radiological findings was evaluated considering histopathological as the gold standard. Performance was evaluated by varying categories considered as TRUE. Across various categories, the range of values were 93.65-96.61% sensitivity, 82.2-97% specificity, 84.44-95.16% PPV, 96.4% NPV, and 89.41-95.71% accuracy (Table 3).

The performance of cytological findings was evaluated considering histopathological as the gold standard. Performance was evaluated by varying categories considered as TRUE. Across various categories, the range of values were 95.83 - 97.60% sensitivity, 90.08-98.33% specificity, 90.37-97.18% PPV, 97.52% NPV, and 93.75-97.40% accuracy (Table 4).
By comparing the data from Tables 3 and 4, it could be inferred that the FNAC offers slightly better results compared to radio-diagnostic tests. These findings were exactly in line with the findings of the correlations presented with the heat map in Figure 2. Heat maps between cytological with histopathological and radiological with histopathological findings in this figure clearly visually outline it.

**Invasive Ductal Carcinoma (IDC) Histopathology Case**

The histopathology was performed for 256 out of 377 cases. After the HE analysis, it was known that one case of IDC histopathology was showing the presence of a tumor area along with necrosis (Figure 3). The cells had a high nuclear-cytoplasmic ratio and presence prominent nucleoli. The section also showed the presence of an atypical mitotic figure. Grade III of IDC was not otherwise specified.

### Table 3. Statistic performance of radiological findings under different characterisation.

<table>
<thead>
<tr>
<th>Statistic Performance</th>
<th>Radiological Finding</th>
<th>I</th>
<th>II</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity (%)</td>
<td>93.65</td>
<td>96.19</td>
<td>96.61</td>
<td></td>
</tr>
<tr>
<td>Specificity (%)</td>
<td>97.00</td>
<td>93.27</td>
<td>82.20</td>
<td></td>
</tr>
<tr>
<td>PPV (%)</td>
<td>95.16</td>
<td>93.52</td>
<td>84.44</td>
<td></td>
</tr>
<tr>
<td>NPV (%)</td>
<td>96.04</td>
<td>96.04</td>
<td>96.04</td>
<td></td>
</tr>
<tr>
<td>Accuracy (%)</td>
<td>95.71</td>
<td>94.74</td>
<td>89.41</td>
<td></td>
</tr>
</tbody>
</table>

*Case category that was considered as true value. I: Only malignant; II: Malignant and suspicious. III: Malignant, suspicious and atypia.

### Table 4. Statistic performance of cytological findings under different characterisation.

<table>
<thead>
<tr>
<th>Statistic Performance</th>
<th>Cytological Finding</th>
<th>I</th>
<th>II</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity (%)</td>
<td>95.83</td>
<td>92.27</td>
<td>97.60</td>
<td></td>
</tr>
<tr>
<td>Specificity (%)</td>
<td>98.33</td>
<td>95.16</td>
<td>90.08</td>
<td></td>
</tr>
<tr>
<td>PPV (%)</td>
<td>97.18</td>
<td>94.69</td>
<td>90.37</td>
<td></td>
</tr>
<tr>
<td>NPV (%)</td>
<td>97.52</td>
<td>97.52</td>
<td>97.52</td>
<td></td>
</tr>
<tr>
<td>Accuracy (%)</td>
<td>97.40</td>
<td>96.15</td>
<td>93.75</td>
<td></td>
</tr>
</tbody>
</table>

*Case category that was considered as true value. I: Only malignant; II: Malignant and suspicious. III: Malignant, suspicious and atypia.
Test of Independence

The chi-square test of Independence between cytological and radiological findings was presented in Table 5. As extremely few patients, with low risk, underwent radiology ultrasound, such cases were ignored from the analysis. The chi-square statistic came out to be 533 at the degree of freedom = 9. The statistic test, \( \chi^2 \), was not in the 95% region of acceptance: \([-\infty: 16.919]\). Hence, it could be inferred that there was a significant dependence between cytological and radiological findings. This finding further corroborated the findings of Figure 1 and 2. In Figure 2, the heat map between cytological and radiological presented a strong correlation between these diagnostic methods.

Discussion

Carcinoma of the breast is the highest morbidity malignancy with fourth ranking for mortality.\(^5,12\) In the present study, the highest percentage of malignant cases is seen in the cases of the eighth decade along with a rising trend from the fifth decade onwards. This is at par with many of previous studies.\(^9,13-15\) The highest number of cases are of fibroadenoma followed by fibrocystic disease of the breast. These findings are similar to the other studies.\(^16-19\)

The alarming increase in case number in the second and third decades is very disturbing and more research is needed in this area. Considering histopathological findings as gold standard\(^20\) and correlating against it maximum false-negative cases were proliferative breast disease with atypia on cytology, code 3 (Yokohama Congress of cytology), which turned out to be florid ductal hyperplasia on histopathological examination.\(^21,22\) The next were cases of fibroadenoma with duct carcinoma in situ that have been missed on FNAC and were easily caught on histopathology, proving that histopathology is a must in all breast lesions.\(^23,24\)

The role of FNAC has been proven as one of the efficient and effective tools for the early diagnosis of benign and malignant cases, especially in the case of breast lesions, along with radiology and clinical correlation. The specificity, PPV, and NPV of this study are comparable to other studies done in the past.\(^17-19,24\)

Radiology also plays a significant role with high sensitivity and accuracy.\(^25,26\) These are equivalent or more in the present study as compared to ones done in the past.

Table 5. Chi-square test of independence between cytological and radiological findings.

<table>
<thead>
<tr>
<th>Cytological Observed</th>
<th>Radiological Observed</th>
<th>Expected Values</th>
<th>Chi-Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Values</td>
<td>2 3 4 5</td>
<td>2 3 4 5</td>
<td>2 3 4 5</td>
</tr>
<tr>
<td>2</td>
<td>103 7 1 1</td>
<td>43.01 16.23 26.38 26.38</td>
<td>83.65 5.25 24.41 24.41</td>
</tr>
<tr>
<td>3</td>
<td>3 26 0 0</td>
<td>11.14 04.20 6.83 6.83</td>
<td>5.95 113.04 6.83 6.83</td>
</tr>
<tr>
<td>4</td>
<td>0 7 51 5</td>
<td>24.20 09.13 14.84 14.84</td>
<td>24.20 0.50 88.14 6.52</td>
</tr>
<tr>
<td>5</td>
<td>0 0 13 59</td>
<td>27.65 10.43 16.96 16.96</td>
<td>27.65 10.43 0.92 104.25</td>
</tr>
</tbody>
</table>
Figure 4. Comparison between (A) Traditional triple test approach and (B) The new proposed alternative approach for developing countries.
This proves that it still indeed plays a significant part in the triad for palpable breast lesions.27,28 Fewer radiology cases, as compared to cytology, are either attributed to the recommendation of the surgeon or the higher cost of the radiological test.

Based on the chi-square test results, it could be concluded that cytolgical and radiological findings are significantly dependent on each other. Also, the tabulated value from the chi-square table, 16.919, is significantly less than that of the calculated value of 533. This implies that the cytological and radiological findings are strongly dependent on each other.

The new proposed approach suggests employing FNAC along with clinical examination as primary diagnosis modality. Radiological examinations will be conducted only when the treating physician needed additional inputs after reviewing the findings of clinical and FNAC examinations. The contrast between the traditional triple-test approach and the new proposed approach is presented in Figure 4.

To establish the findings of this study, a wider and multi-center study is required. Additionally, as only the retrospective samples were processed as a part of this study, many cancer-causing parameters such as family history, and repeated complaints of breast lumps, etc. were not included in the study. A prospective study in regard can bring forth fresh insights. It will also enable us to see the real-time benefits of the proposed new alternative approach when compared with the traditional triple-test approach.

Conclusion

With the strong correlation between cytolgical and radiological findings obtained in this study along with superior results of cytological diagnosis over radiological, we can infer that a combination of FNAC with clinical examination can be used in place of triple test to make the diagnosis of the breast lesions and chart down the treatment without making significant compromises on the risk of incorrect treatment. If deemed necessary, radiological tests can be advised subsequently. This way, a significant reduction in the workload of radiologists, reduction in overall treatment cost, and equity of health can be achieved.

Author Contribution

PS and MMJ were involved in the conception and planning of the research, PS and CD performed the data acquisition/collection, BB, MMJ and PS calculated the experimental data and performed the analysis, PS drafted the manuscript, BB designed the figures, and CD and PS aided in interpreting the results. PS, MMJ, CD, and BB took part in giving critical revision of the manuscript.

References