Original Research Paper

Volume - 15 | Issue - 06 | June - 2025 | PRINT ISSN No. 2249 - 555X | DOI : 10.36106/ijar

Oral Pathology

COMPARATIVE ANALYSIS OF AUTOMATED AND PATHOLOGIST VISUAL SCORING OF IMMUNOHISTOCHEMISTRY IMAGES

Dr. R Rajeev*	Associate Professor, Government Dental College, Thiruvananthapuram *Corresponding Author
Dr. Parul Dixit	Junior Resident, Government Dental College, Thiruvananthapuram
Dr. Lekshmi G	Former Junior Resident, Government Dental College, Thiruvananthapuram
Dr. S.K Padmakumar	HOD & Professor, Government Dental College, Thiruvananthapuram

ABSTRACT

Introduction: Immunohistochemistry (IHC) is a cornerstone in diagnostic pathology, especially for cancer, aiding in the detection and quantification of molecular markers in tissue samples. Traditional manual scoring methods, though widely practiced, are prone to observer bias and are time-intensive. With advancements in digital imaging, automated platforms like ImageJ offer a potential alternative for objective and efficient IHC analysis. **Methods:** This study evaluated 72 histopathologically confirmed cases comprising oral epithelial dysplasia (OED) and oral squamous cell carcinoma (OSCC). ALDH1A1 expression was assessed via IHC, and tissue slides were scored manually by pathologists and digitally using the ImageJ software. Three metrics were compared: intensity, quantity, and Immunoreactive Score (IRS). The correlation between manual and automated methods was analyzed using Spearman's correlation coefficient and Cohen's Weighted Kappa. **Results:** Findings revealed a high concordance between manual and ImageJ-based scoring. Intensity scores matched perfectly for OED and showed strong agreement in OSCC. Quantity scores exhibited slight variation, affecting IRS scores. Spearman's correlation demonstrated significant positive relationships across all parameters ($p < 0.001$), while Cohen's Kappa indicated moderate to strong agreement (κ ranging from 0.637 to 1.00). **Conclusion:** This study highlights the effectiveness of an ImageJ-compatible open-source plugin for automated IHC scoring. The method showed strong concordance with manual analysis, supporting its potential for rapid, objective, and scalable evaluation of biomarker expression in clinical and research settings.

KEYWORDS : Immunohistochemistry, ImageJ, Automated Scoring, Oral Squamous Cell Carcinoma

1. INTRODUCTION

Immunohistochemistry (IHC) is a cornerstone in diagnostic pathology, particularly for cancer detection and classification. It enables the visualization and quantification of molecular markers in formalin-fixed paraffin-embedded (FFPE) tissue sections using antibody-based staining. Traditionally, these assays rely on manual scoring by pathologists who evaluate staining intensity and distribution patterns visually. Despite its widespread use, this method has inherent limitations. It is time-consuming, subjective, and often restricted to ordinal data, leading to potential inter-observer variability and decreased reproducibility[1]. These limitations become particularly problematic when attempting to detect subtle changes in biomarker expression that may have significant clinical implications[2].

In recent years, advancements in digital pathology and image analysis have prompted the development of computational tools aimed at enhancing the accuracy and consistency of IHC evaluations[3]. One such tool is ImageJ, an open-source, Java-based image processing program that aids in quantitative immunoscoring, allowing for more objective, reproducible, and scalable analysis of stained tissue images[4]. This study investigates the efficacy of an ImageJ-based automated scoring method by comparing its results to conventional manual scoring performed by experienced oral pathologists. The primary focus is on assessing ALDH1A1 expression, a known cancer stem cell marker, in cases of oral epithelial dysplasia (OED) and oral squamous cell carcinoma (OSCC)[5]. By validating the concordance between automated and manual methods, this research aims to support the integration of digital tools into routine pathology practice for enhanced diagnostic precision and workflow efficiency.

2. MATERIALS AND METHODS

2.1 Case Selection

This study included a total of 72 formalin-fixed, paraffin-embedded (FFPE) tissue blocks, histopathologically diagnosed as either oral epithelial dysplasia or oral squamous cell carcinoma. The sample distribution comprised 12 cases each of mild, moderate, and severe epithelial dysplasia, and well, moderately, and poorly differentiated OSCC. All cases were selected based on confirmed histological criteria and further analysed.

2.2 Immunohistochemistry Protocol

Tissue sections of 4 µm thickness were cut and mounted on poly-L-lysine-coated slides. Sections were deparaffinized in xylene and rehydrated through graded alcohols. Endogenous peroxidase activity was blocked using 3% hydrogen peroxide. Antigen retrieval was performed using a pressure cooker in citrate buffer (pH 6.0). After cooling, slides were incubated overnight at 4°C with rabbit monoclonal anti-ALDH1A1 primary antibody (1:400 dilution in PBS). Secondary antibody conjugated with horseradish peroxidase (HRP) was applied, followed by chromogenic detection using 3,3'-diaminobenzidine (DAB). Slides were counterstained with hematoxylin, dehydrated, and mounted for microscopic evaluation. Negative controls were processed in parallel by omitting the primary antibody.

2.3 Visual Scoring Method

Two independent oral pathologists evaluated the slides using light microscopy. Scoring was based on two parameters - Intensity Score and Quantity Score and Immunoreactive Score (IRS). The IRS was calculated as the product of intensity and quantity, following this scoring criteria[6](Table 1):

Table-1: Scoring Criteria for Immunohistochemical Evaluation.

Quantity Score	Intensity Score	IRS Score
0 = no positive cells	0 = No colour	0.1 = Negative
1 = < 10% positive cells	1 = Mild	2.3 = Mild
2 = 10.50% positive cells	2 = Moderate	4.8 = Moderate
3 = 51.80% positive cells	3 = Intense	9.12 = Strongly Positive
4 = ≥ 80% positive cells	–	–

2.4 Image Acquisition and Software Analysis

Representative fields were captured at 10x magnification using a Jenoptik ProgResC3 camera-mounted microscope. All images were standardized for white balance and color density. JPEG-formatted images were processed using ImageJ software bundled with Java 8. Image analysis involved thresholding to define positive staining areas, from which the software extracted mean intensity, percentage of positive area, and derived the IRS(Fig-1)

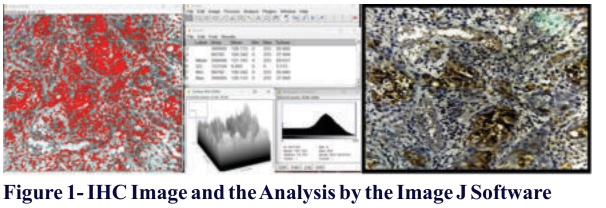


Figure 1- IHC Image and the Analysis by the Image J Software

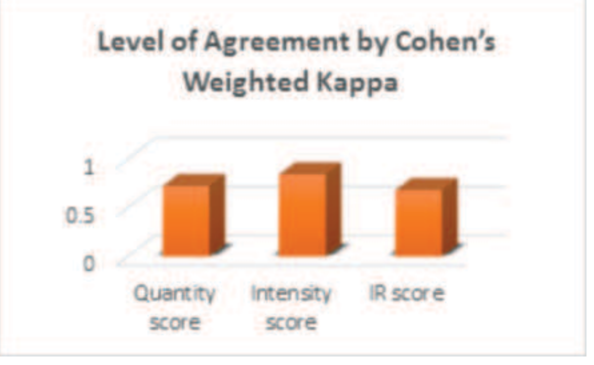
Following criteria is used for scoring.[7] (Table 2)

Table 2: Scoring Criteria for Immunohistochemical Analysis by Image J

Intensity score	Quantity Score	IRS Score
100-130 - 1	Less than 10% - 1	0,1=Negative
130-160 - 2	10- 50 % - 2	2,3=Mild
160-180 - 3	51- 80 - 3	4,8=Moderate
	More than 80- 4	9,12=Strongly Positive

3. RESULTS

The comparison between software and manual scoring methods for oral epithelial dysplasia (OED) and oral squamous cell carcinoma (OSCC) cases showed strong concordance in intensity scores (OED: Cohen's Weighted Kappa $\kappa = 1.00$, Spearman's rho $\rho = 1.00$; OSCC: $\kappa = 0.719$, $\rho = 0.582$; $p < 0.001$) and moderate agreement for quantity (OED: $\kappa = 0.687$, $\rho = 0.729$; OSCC: $\kappa = 0.640$, $\rho = 0.686$) and immunoreactive scores (IRS) (OED: $\kappa = 0.637$, $\rho = 0.639$; OSCC: $\kappa = 0.668$, $\rho = 0.514$)[8](Graph 1)



Graph 1-Level of Agreement Between Software and Manual Method in Quantity, Intensity and Immunoreactive Score.

4. DISCUSSION

The findings of this study demonstrate strong concordance between automated software-based and manual scoring methods for immunohistochemical (IHC) evaluation in oral epithelial dysplasia (OED) and oral squamous cell carcinoma (OSCC), particularly in assessing staining intensity. Moderate agreement was observed for both quantity and Immunoreactive Scores (IRS), as pathologists perform manual counting of individual cells, whereas the software relies on percentage-based estimations. This discrepancy can be mitigated by incorporating a variety of plugins, providing scope to enhance the software's analytical capabilities[8]. These noteworthy results, in agreement with Smith et al., suggest that either method can be chosen based on resource availability without sacrificing accuracy, diagnostic reliability, or reproducibility[9]. The high to moderate correlations obtained via Spearman's rho and Cohen's Weighted Kappa further support the reliability and consistency of software-based quantification compared to traditional manual evaluation. The modular nature of platforms like ImageJ allows integration of various plugin tools, which can significantly improve scoring accuracy and reproducibility[1]. Such automated systems offer a promising solution to the challenges of inter-observer variability often encountered in manual scoring methods[10].

By integrating automated image analysis, pathology laboratories and research institutions can optimize resource utilization, reduce time and labour demands, and enhance data reproducibility and objectivity[3]. Nonetheless, this study acknowledges certain limitations, including a relatively small sample size and analysis restricted to a single biomarker. Future research should aim to validate these findings across larger, more heterogeneous cohorts and a wider range of biomarkers. As both staining protocols and digital analysis tools continue to evolve, periodic re-evaluation of automated methods will be essential to ensure sustained accuracy and clinical relevance[10].

5. CONCLUSION

Evidence from this study supports the conclusion that automated software-based scoring for immunohistochemical analysis in oral epithelial dysplasia and oral squamous cell carcinoma shows strong agreement with manual evaluation, particularly for intensity scores, and moderate agreement for quantity and immunoreactive scores. The validated reliability and consistency of the software suggest it is a valuable tool that can enhance diagnostic accuracy while optimising resources and reducing observer variability. Despite some limitations, integrating automated image analysis into routine pathology practice holds significant promise.

REFERENCES

1. Varghese, F., Bukhari, A. B., Malhotra, R., & De, A. (2014). IHC Profiler: An open source plugin for the quantitative evaluation and automated scoring of immunohistochemistry images of human tissue samples. PLOS ONE, 9(5), e96801. https://doi.org/10.1371/journal.pone.0096801
2. Taylor, C. R., & Levenson, R. M. (2006). Quantification of immunohistochemistry—Issues concerning methods, utility and semiquantitative assessment II. Histopathology, 49(4), 411–424. https://doi.org/10.1111/j.1365-2559.2006.02513.x
3. Cizkova, K., Foltynkova, T., Gachechiladze, M., & Tauber, Z. (2021). Comparative analysis of immunohistochemical staining intensity determined by light microscopy, ImageJ and QuPath in placental Hofbauer cells. Acta Histochemica et Cytochemica, 54(1), 21–29. https://doi.org/10.1267/ahc.20-00033
4. Hartig, S. M. (2013). Basic image analysis and manipulation in ImageJ. Current Protocols in Molecular Biology, 102(1), 14-5. https://doi.org/10.1002/0471142727.mb1415s102
5. Gupta, V., Maurya, M. K., Agarwal, P., Kumar, M., Sagar, M., Raghuvanshi, S., & Gupta, S. (2022). Expression of aldehyde dehydrogenase 1A1 in oral squamous cell carcinoma and its correlation with clinicopathological parameters. National Journal of Maxillofacial Surgery, 13(2), 208-215.
6. Yaghjyan, L., Heng, Y. J., Baker, G. M., Bret-Mounet, V., Murthy, D., Mahoney, M. B., ... & Tamimi, R. M. (2022). Reliability of CD44, CD24, and ALDH1A1 immunohistochemical staining: Pathologist assessment compared to quantitative image analysis. Frontiers in Medicine, 9, 1040061.
7. Guirado, R., Carceller, H., Castillo-Gomez, E., Castren, E., & Nacher, J. (2018). Automated analysis of images for molecular quantification in immunohistochemistry. Heliyon, 4(6).
8. McKinney, K. (2018). Quantitative immunohistochemistry using ImageJ. Journal of Visualized Experiments, (136), e57623. https://doi.org/10.3791/57623
9. Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. Biometrics, 33(1), 159–174. https://doi.org/10.2307/2529310
10. Smith, M. A., Jones, T., Brown, L., & Patel, N. (2023). Digital pathology and AI-based scoring of biomarkers in oncology: A validation study. Modern Pathology, 36(2), 200–212. https://doi.org/10.1038/s41379-022-01110-w