

## Original Research Article

### **The role of cell detection tool of QuPath-0.3.2 software in differentiating benign cervical epithelium from malignant epithelial clusters.**

#### **ABSTRACT**

**Aim:** The study aims to be an initial assessment of the ability of the Qu Path software to quantitatively distinguish between benign cervical epithelium and non-keratinizing squamous cell carcinoma cell clusters.

**Study design:** We carried out a retrospective cross-sectional study of nineteen slides of cervical biopsies of non-keratinizing squamous cell carcinoma(NKSCC) of the cervix.

**Place and Duration of Study:** These are from cases of cervical carcinomas seen in UniOsun Teaching Hospital, Osogbo between 1<sup>st</sup> January 2006 and 31<sup>st</sup>December2015.

**Materials and Methods:** These slides had been prepared by standard laboratory protocol and stained with hematoxylin and eosin. The slides were scanned using CS2 APERIO digital slide scanner with standard settings.

The scanned slides were viewed using QuPath-0.3.2 software. For each normal or NKSCC cell found in the randomly sampled cluster of the biopsy, the nuclear and cytoplasmic morphometric features were analyzed using Microsoft Excel and SPSS version 20.

**Results:** The cells in the normal clusters were 13,774 and that of the NKSCC was 26,114. There is a higher mean nuclear area in NKSCC cells compared to normal cells. **The malignant cells** have a higher mean nuclear area, circularity, and perimeter than cells of the benign epithelium. There was a higher nuclear hematoxylin OD (optic density) mean, hematoxylin OD max, and hematoxylin OD sum in the malignant cells. There was a higher cell eccentricity, mean cytoplasmic eosin OD, maximum eosin OD and minimum eosin

OD in the malignant cells. The maximum nuclear caliber is higher in cells of NKSCC than that of benign epithelium.

**Conclusion:** The QuPath software (version 0.3.2) can be a useful aid to differentiate non-keratinizing squamous cell carcinoma(NKSCC) from normal squamous cervical epithelium especially in fragmented biopsies.

## **KEYWORDS**

Cervix Uteri, Carcinoma, Squamous Cell, Software, Computers

## **INTRODUCTION**

Many researchers have tried to use image analysis to improve tissue diagnosis of cancer. [1-4] There has been a lot of research that showed the possible role of computer-aided cell morphometry in diagnosis of lesions and predicting clinical outcomes[5-8]. Most of these works studied nuclear morphometry in terms of nuclear size and area. The QuPath-0.3.2 software is a relatively new image analysis software that has been used by many researchers in evaluating scanned histopathology slides. It is widely used because its open-source and flexible in adapting to various imaging requirements [9]. Apart from measuring nuclear and cytoplasmic dimensions, it can measure intensity of haematoxylin and eosin stains. Hence, under standard conditions, it can give a measure of hyper-chromaticity and can be used to measure the amount of keratin in the cytoplasm.

It is unique because at appropriate settings, it can measure cell parameters that were usually evaluated subjectively. It has introduced objectivity in parameters like cell size, hyperchromatism, nuclear borders and size variation It is capable of objectively assessing the

perimeter and circularity of nuclei and cell membranes, optical densities of hematoxylin and eosin in nucleus and cytoplasm and many other parameters.

Cervical cancer is one of the most common cancers in women [10-12]. Diagnosis of cervical cancers is usually very easy for histopathologists but from our experience there are situations where the epithelial tissues are fragmented in biopsies, and it can sometimes be challenging to differentiate a benign epithelial fragment from a malignant one.

The study aims to be an initial assessment of the ability of the Qu Path software to quantitatively distinguish between benign cervical epithelium and cell clusters of a non-keratinizing squamous cell carcinoma. In this study, we want to see if there are differences between the measurements of cellular parameters of benign cervical epithelium and that of NKSCC cell clusters.

### **Materials and Methods.**

We carried out a retrospective study of all eighty-seven cases of cervical biopsies of non-keratinizing squamous cell carcinoma(NKSCC) of the cervix from cases in UniOsun Teaching Hospital, Osogbo between 1<sup>st</sup> January 2006 and December 31<sup>st</sup>, 2015. The histological slides of the cervical biopsies had been prepared by standard laboratory protocol and stained with hematoxylin and eosin. All slides that are adequately processed, properly stained with hematoxylin and eosin and without widespread artefacts that could interfere with the image analysis process were included in the study. All poorly stained slides or slides without adequate tumor tissue or with widespread artefacts were excluded from the study. Slides for nineteen biopsies met the inclusion criteria and were scanned using CS2 APERIO digital slide scanner with standard settings.

The scanned slides were viewed using QuPath-0.3.2 software. The clusters were annotated using the wand and brush tool of QuPath software. The full thickness of the normal

epithelium seen in the biopsies and malignant epithelial tumors were annotated. Areas of cervical intraepithelial neoplasia, or with artefacts were not annotated for analysis. We selected four areas of normal cervical epithelium seen in the biopsies and two or three clusters of NKSCC in each case for analysis. The number of clusters were randomly selected and were limited to maximum of three clusters to limit the size of the data to prevent processing problems with the software used. The setup parameters of the cell detection tool of the QuPath software were adjusted to optimally detect the cell structures. The nucleus parameters were set at 8 $\mu\text{m}$  background radius, 0  $\mu\text{m}$  median filter radius, 1.5 $\mu\text{m}$  Sigma, 10 $\mu\text{m}^2$  minimum area and 400 $\mu\text{m}^2$  maximum area. The intensity parameters were set at threshold 0.1 and maximum background intensity at 2. It was set to split by shape option. The cell parameter was set to 5 $\mu\text{m}$  cell expansion to include the nucleus. The general parameters were set to smooth boundaries and make measurements.

For each normal or NKSCC cell, the following parameters were measured; nuclear area, nuclear perimeter, nuclear circularity, hematoxylin OD mean, haematoxylin OD sum, maximum hematoxylin OD, nuclear perimeter, maximum nuclear caliper, cell size (area), cell perimeter, maximum caliper of the cell, cell eccentricity, cytoplasmic eosin OD, eosin OD mean, eosin OD std dev, eosin OD max, and nucleocytoplasmic ratio.

The result of the analysis was exported to Microsoft excel and further analyzed using Statistical Package for Social Sciences (SPSS version 20).

## **Results**

The cells in the normal clusters were 13,774 and that of the **NKSCC** was 26,114. The result of the analysis of the nuclear parameters for each of the normal tumour cells and malignant tumour cells are shown in Table 1. The result of the analysis of the general cell and cytoplasmic parameters for each of the normal tumour cells and malignant tumour cells are shown in Table 2. Figures 1 and 2 are photomicrographs showing a typical focus of a scanned

image and how the software measures the colors and cell morphometrics. Figures 3 to 10 shows the box plot of selected measurements to show how the values of that of the normal cells compare with that of the cells of non-keratinizing squamous cell carcinoma.

Table 1: Nucleus morphometry of normal versus non-keratinizing squamous cell carcinoma (NKSCC) cells by QuPath software

		Mean	Minimum	Maximum	Standard Deviation
Nucleus: Area ( $\mu\text{m}^2$ )	Normal	25.863908	8.1446	393.2338	22.6771056
	NKSCC	29.012483	8.1446	337.2393	22.9758059
Nucleus: Perimeter ( $\mu\text{m}$ )	Normal	22.789072	10.9432	317.8154	13.5919702
	NKSCC	24.132938	10.9082	250.6095	14.4837531
Nucleus: Circularity	Normal	.643234	.0424	1.0000	.1652224
	NKSCC	.667825	.0675	1.0000	.1790836
Nucleus: Max caliper ( $\mu\text{m}$ )	Normal	8.605289	3.5707	84.6504	4.4883323
	NKSCC	8.922111	3.5667	67.3739	4.4036294
Nucleus: Hematoxylin OD mean	Normal	-.071605	-.3430	.8565	.0688747
	NKSCC	.206656	-.3350	1.0557	.1975921
Nucleus: Hematoxylin OD sum	Normal	-8.184069	-199.7528	374.3051	13.0300004
	NKSCC	24.623945	-209.0219	814.0132	41.1415668
Nucleus: Hematoxylin OD max	Normal	.054509	-.2363	1.4449	.1389325
	NKSCC	.404171	-.2590	2.9058	.2846170

Table 2: General cell and cytoplasm morphometry of normal versus non-keratinizing squamous cell carcinoma (NKSCC) cells by QuPath software

		Mean	Minimum	Maximum	Standard Deviation
Cell: Area ( $\mu\text{m}^2$ )	Normal	94.806317	18.0709	1063.1311	62.9829983
	NKSCC	98.208803	17.5619	935.6165	60.1590034
Cell: Perimeter ( $\mu\text{m}$ )	Normal	37.812425	16.6072	210.9153	12.4737122
	NKSCC	38.765443	15.7195	181.5948	12.3680223
Cell: Max caliper( $\mu\text{m}$ )	Normal	14.034288	5.9618	92.1649	4.8840273
	NKSCC	14.415414	5.5595	71.7298	4.7107905
Cell: Eccentricity	Normal	.717609	.0703	.9878	.1476829
	NKSCC	.725056	.0266	.9870	.1426850
Cytoplasm: Eosin OD mean	Normal	.389600	.0323	1.1602	.1090999
	NKSCC	.585789	.0694	1.0809	.1551859
Cytoplasm: Eosin OD std dev	Normal	.137770	.0152	.3411	.0455560
	NKSCC	.140993	.0164	.4692	.0587658
Cytoplasm: Eosin OD max	Normal	.702618	.1342	1.6967	.1392055
	NKSCC	.940380	.3362	3.1531	.2324998
Nucleus/Cell area ratio	Normal	.277967	.0558	.7496	.1031667
	NKSCC	.290834	.0622	.7224	.0891329

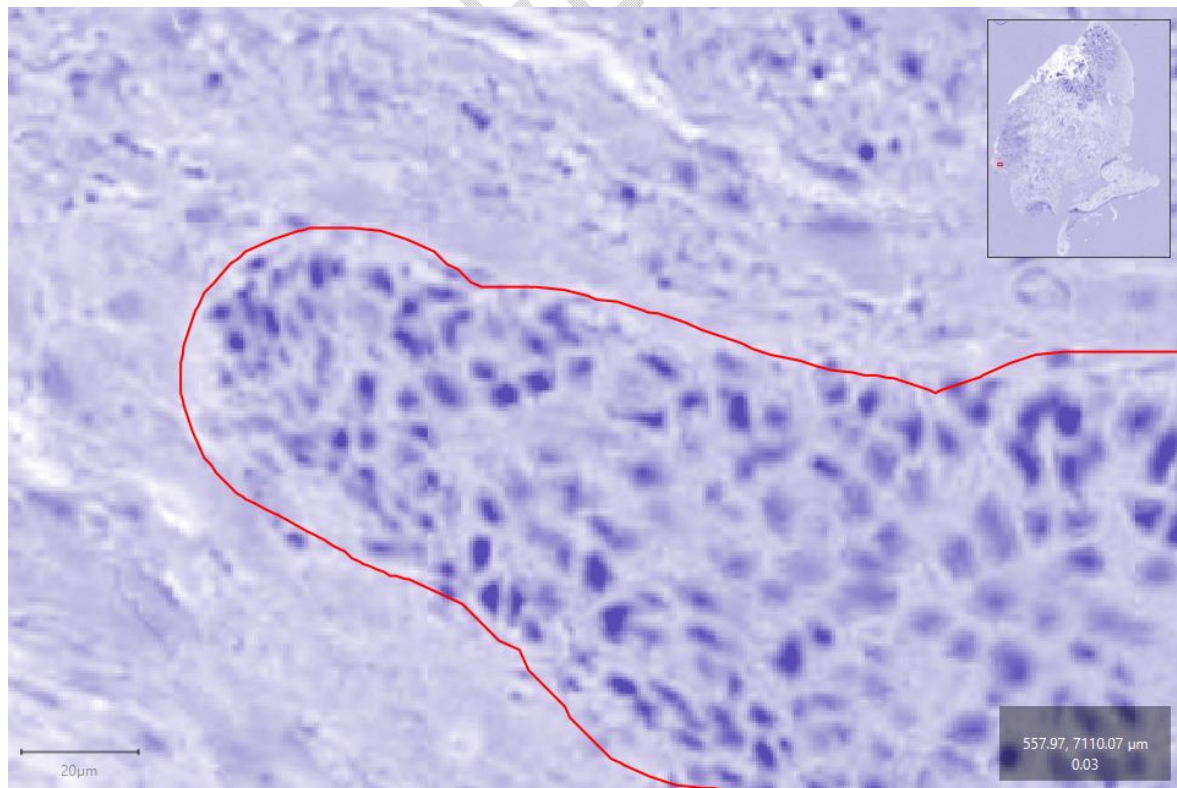


Figure 1. Photomicrograph of a cluster of cancer cells (hematoxylin only). The

photomicrograph shows blue color view to represent the hematoxylin stain. The optic density of the color is measured by the software. Other colors have been digitally excluded by the QuPath software.

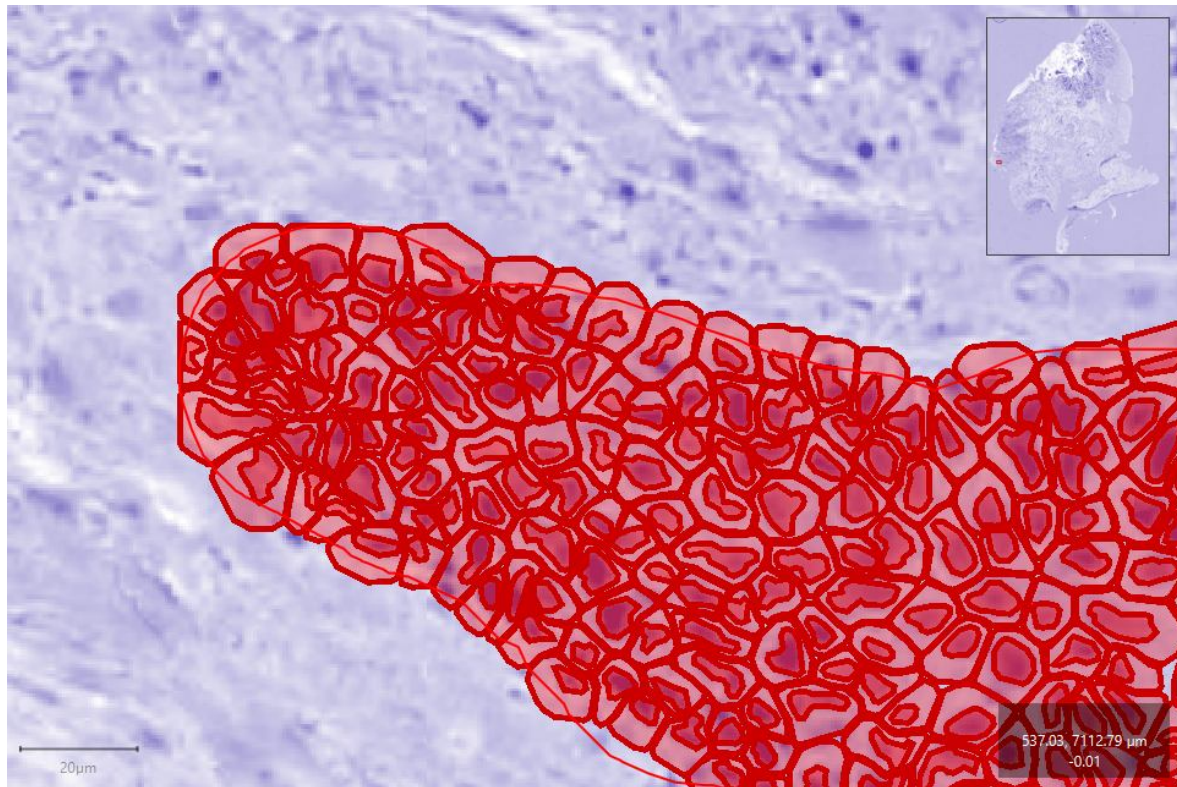


Figure 2 Photomicrograph of nuclei and cytoplasm morphometry by the software. This photomicrograph shows cell detection approximation by the QuPath software. This is used to estimate the size of the nuclei and cytoplasm.

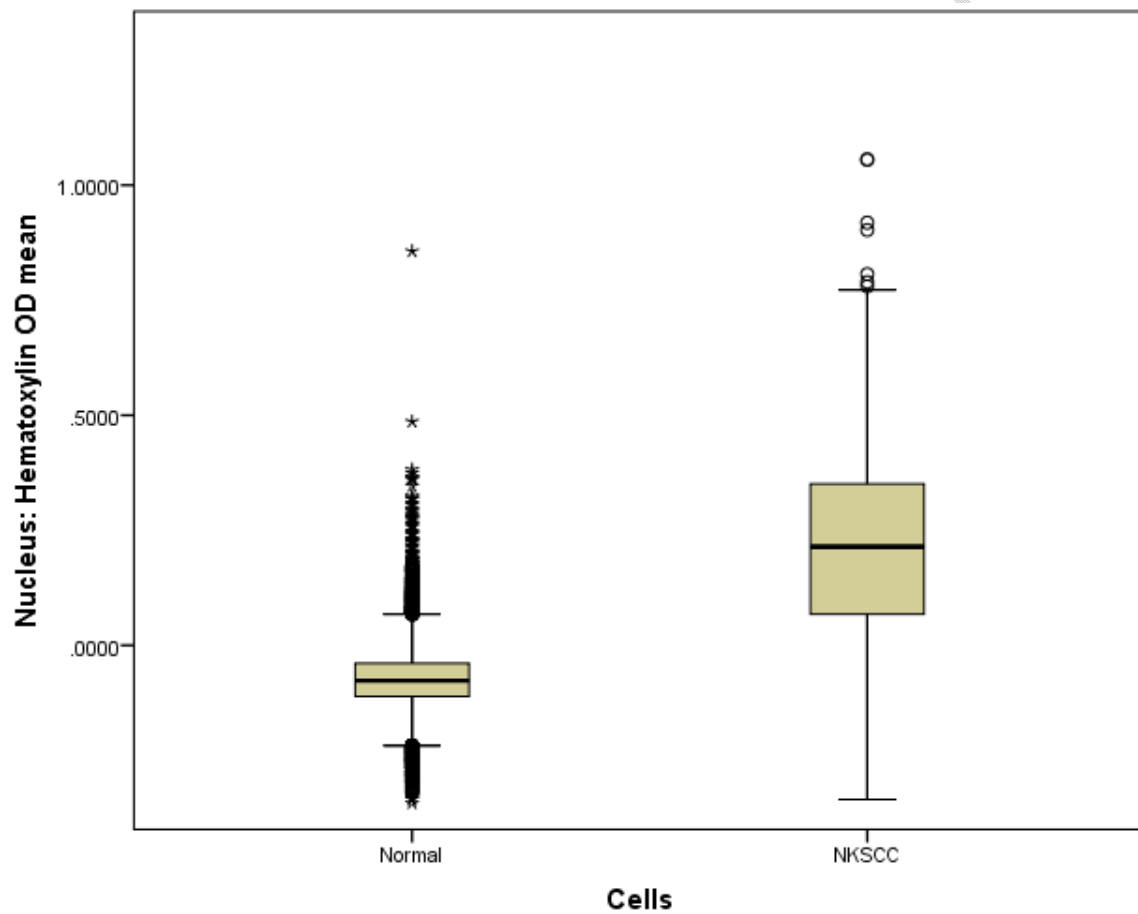


Figure 3. Boxplot of nucleus hematoxylin OD mean of normal and non-keratinizing squamous carcinoma cells. The box-plot above shows that most malignant cells have higher nuclei hematoxylin OD mean than the normal cells. ( $p < 0.001$ )

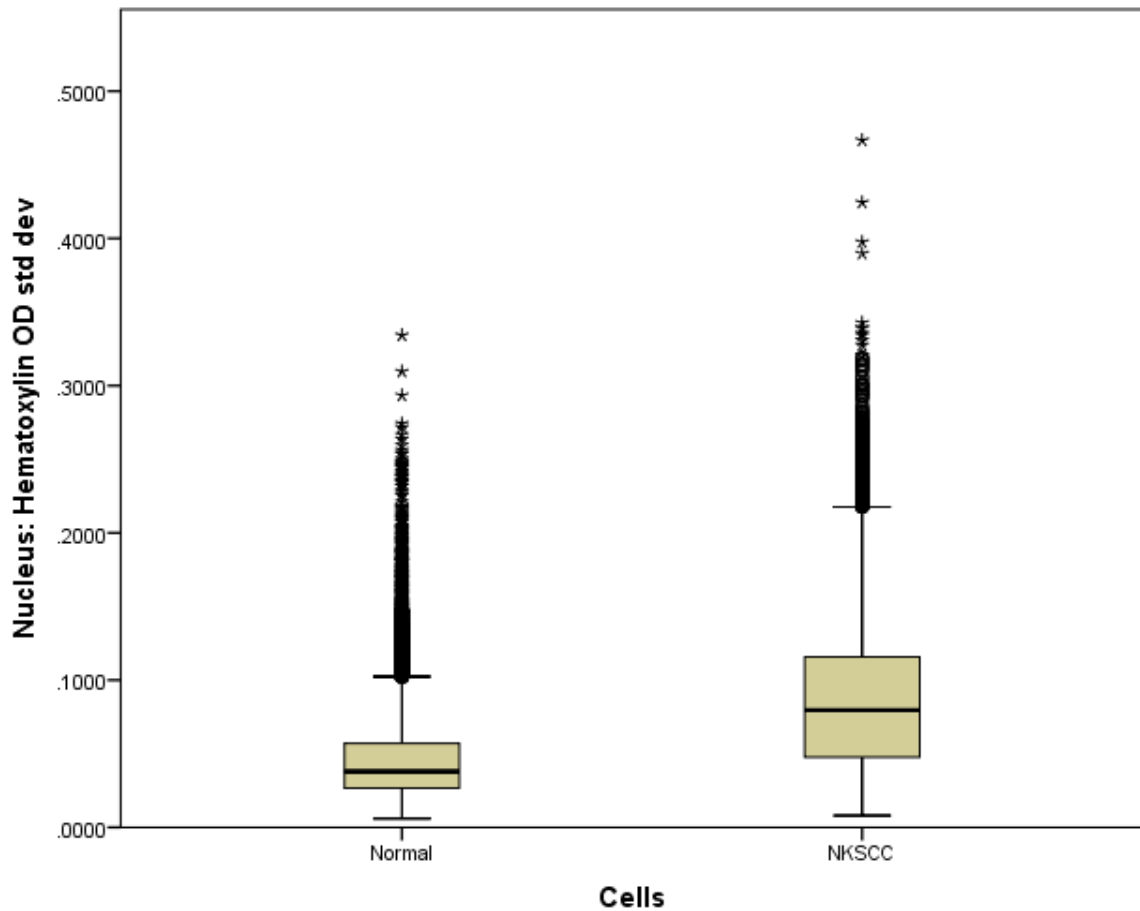


Figure 4. Boxplot of nucleus hematoxylin OD standard deviation for normal and non-keratinizing squamous carcinoma cells. The box-plot above shows that most malignant cells have higher nuclei hematoxylin OD standard deviation than the normal cells. ( $p < 0.001$ )

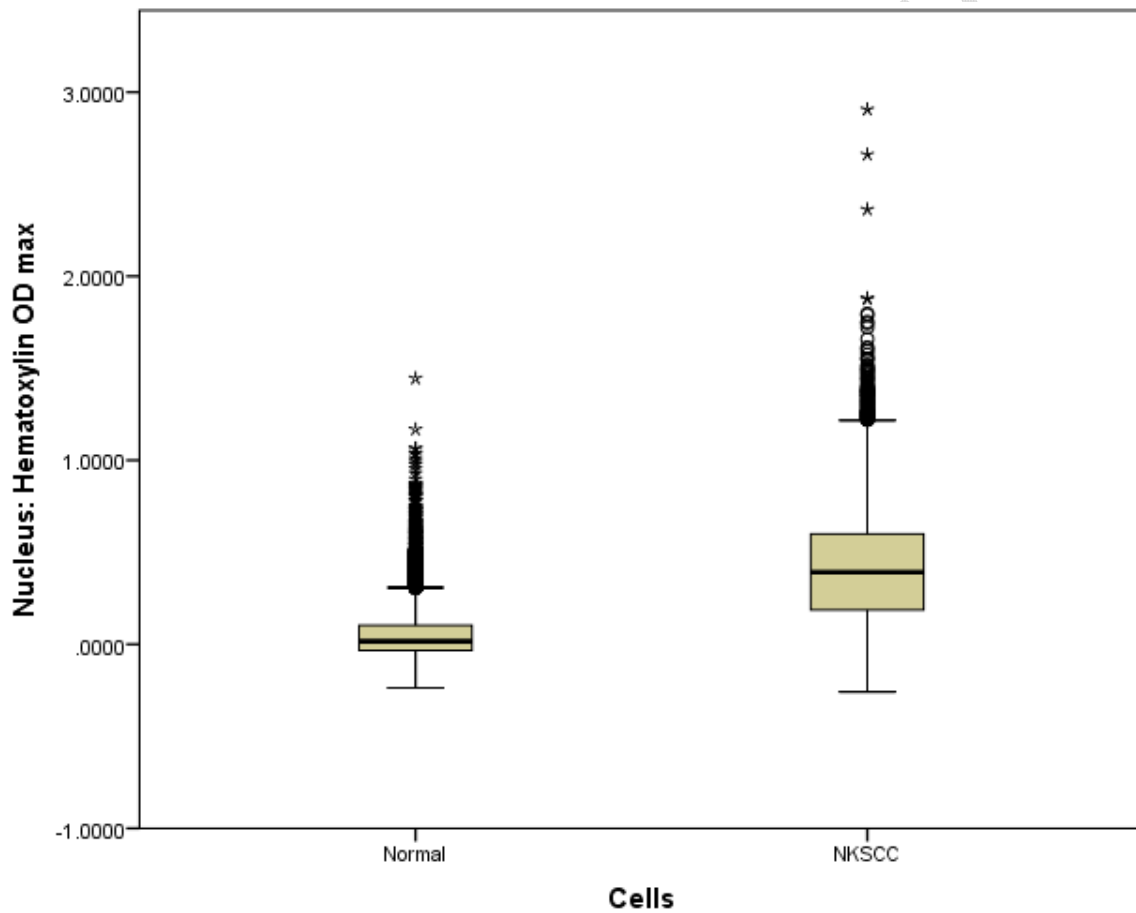


Figure 5. Boxplot of maximum nucleus hematoxylin OD for normal and non-keratinizing squamous carcinoma cells. The box-plot above shows that most malignant cells have higher maximum (max) nuclei hematoxylin OD than the normal cells. ( $p < 0.001$ )

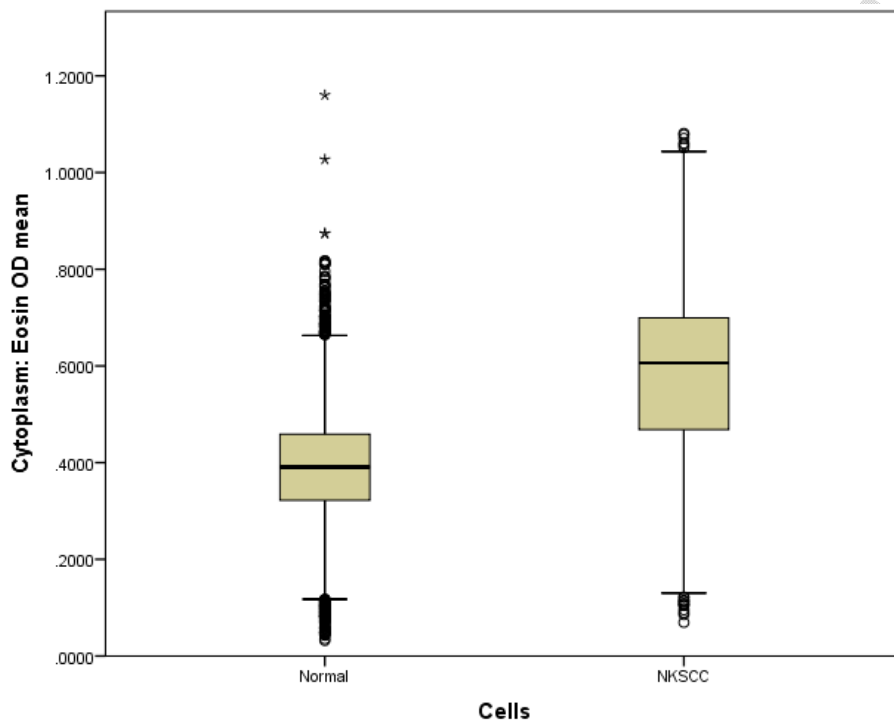


Figure 6. Boxplot of mean cytoplasmic eosin OD for normal and non-keratinizing squamous carcinoma cells. The box-plot above shows that most malignant cells have higher cytoplasmic eosin OD mean than the normal cells. ( $p < 0.001$ )

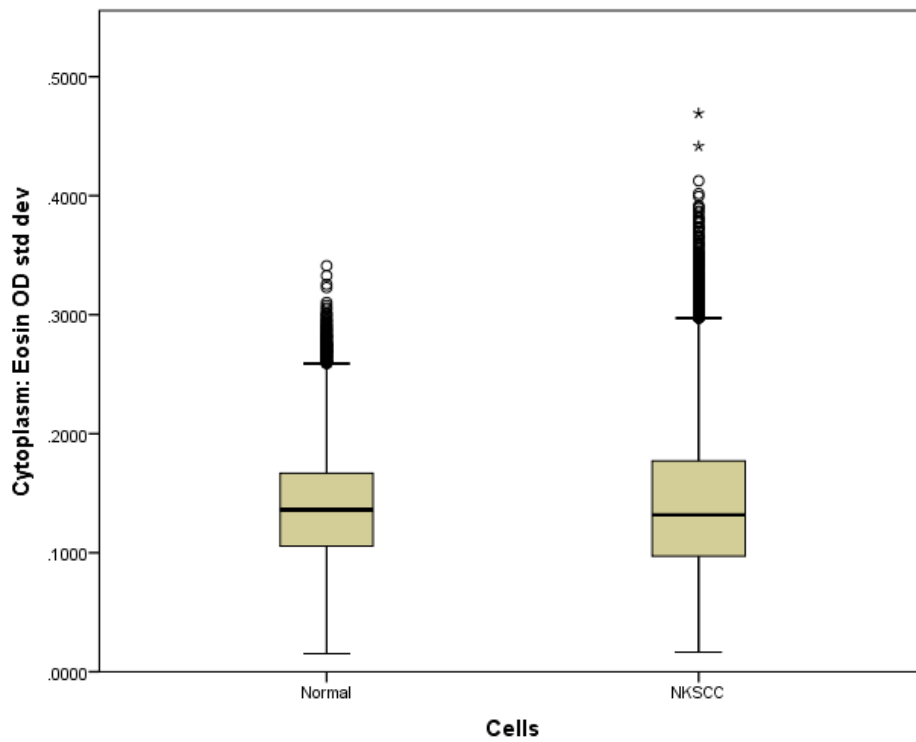


Figure 7. Boxplot of cytoplasmic eosin OD standard deviation for normal and non-keratinizing squamous carcinoma cells. The box-plot above shows that most malignant cells have similar cytoplasmic eosin OD standard deviation (std dev) with the normal cells.

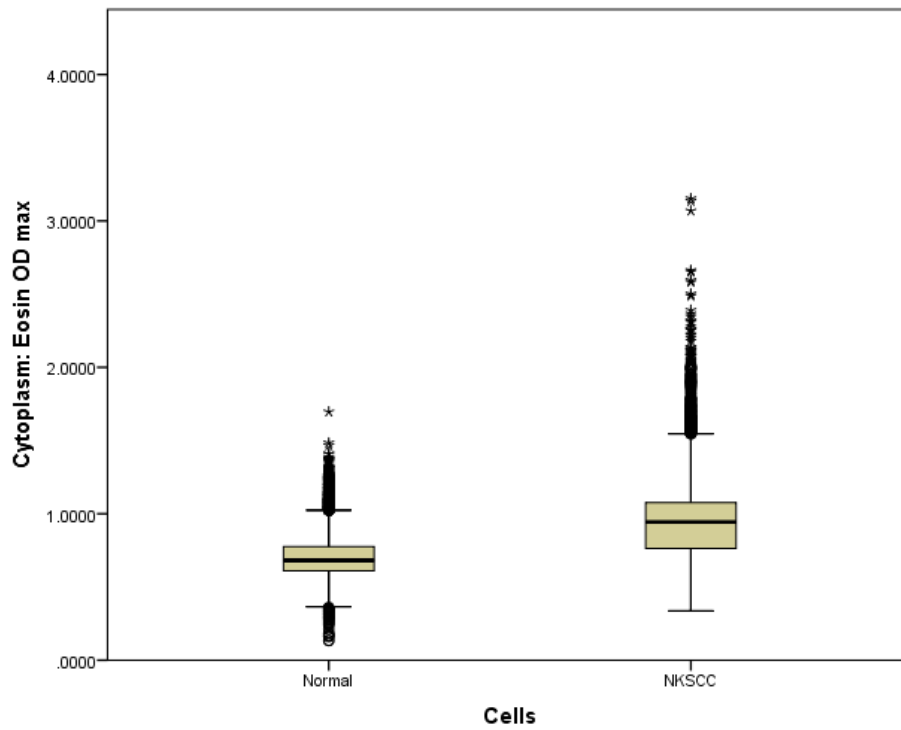


Figure 8. Boxplot of maximum cytoplasmic eosin OD for normal and non-keratinizing squamous carcinoma cells. The box-plot above shows that most malignant cells have higher maximum (max) cytoplasmic eosin OD. ( $p < 0.001$ )

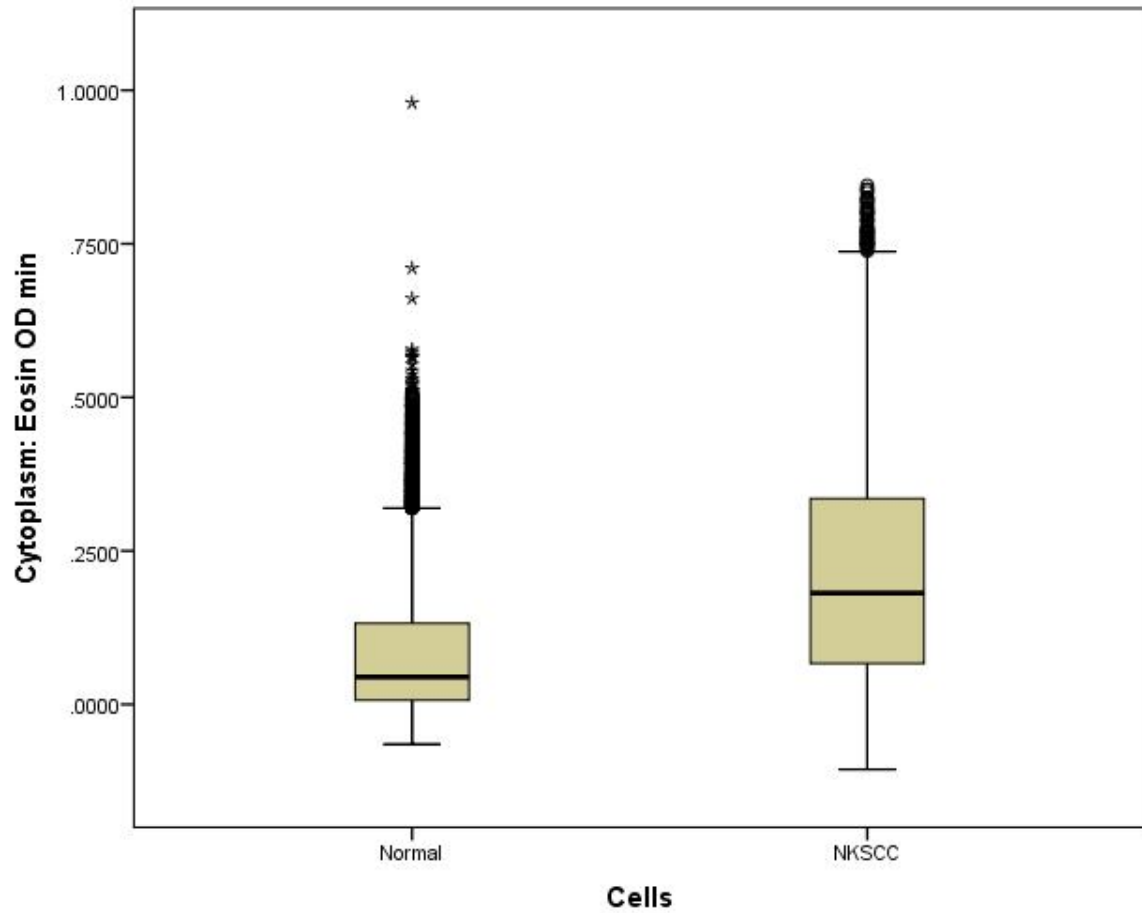


Figure 9. Boxplot of minimum cytoplasmic eosin OD for normal and non-keratinizing squamous carcinoma cells. The box-plot above shows that most malignant cells have higher minimum cytoplasmic eosin OD. ( $p < 0.001$ )

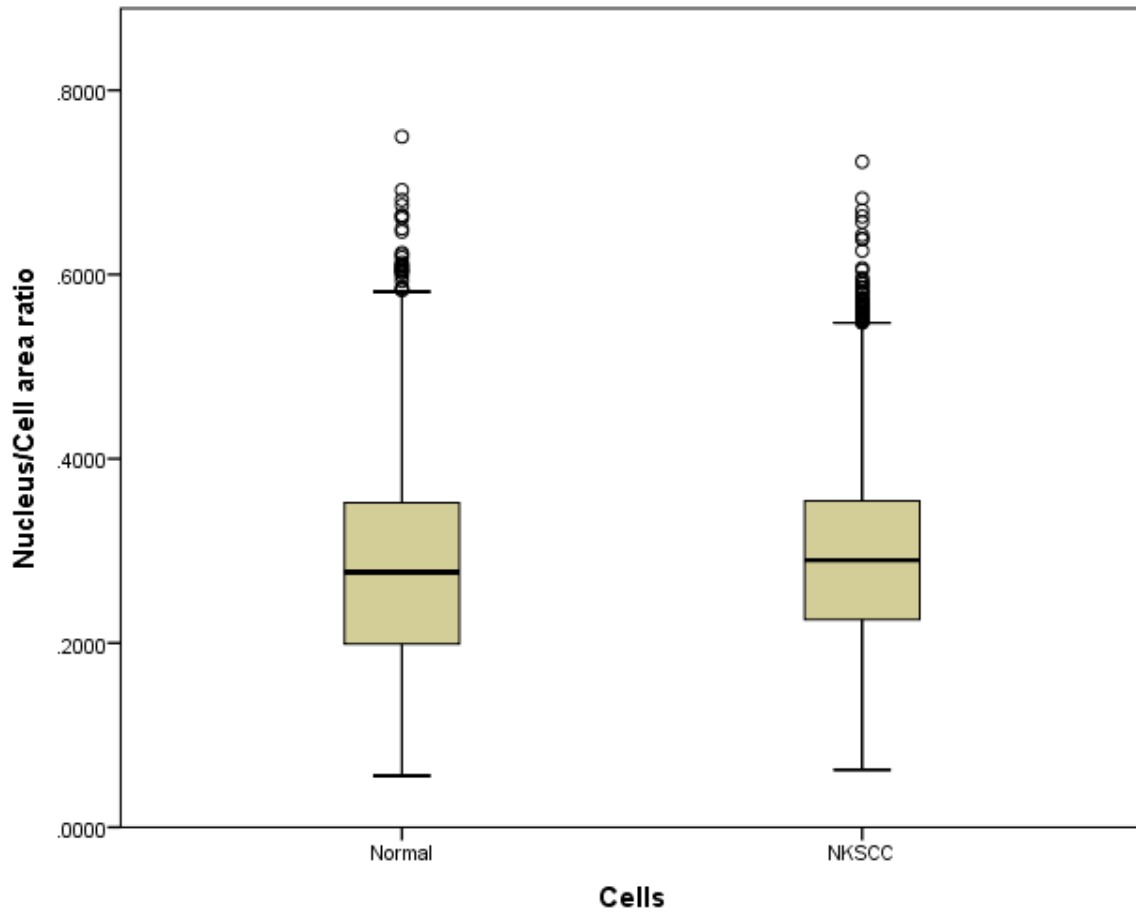


Figure 10. Boxplot of nucleus-cell ratio of normal and non-keratinizing squamous carcinoma cells. The box-plot above shows that most malignant and normal cells have similar nucleocytoplasmic ratio measured by the software.

## **Discussion**

The results show clear differences between the cells of the normal cervical epithelium and non-keratinizing squamous cell carcinoma (NKSCC).

Malignant cervical cells are known to have larger nuclei [13] which is expected to be shown as larger nuclear area. This was supported by findings in our study (Table 1) which revealed a higher mean nuclear area in non-keratinizing squamous cell carcinoma cells compared to normal cells. Although the maximum nuclear area was noticed to be higher in the normal cell, the measurement is likely to be an outlier and may be due to imperfect estimation by the software.

Nuclear perimeter also shows remarkable difference between cells of the normal epithelium and NKSCC. The cells of the malignant cells have a much higher mean nuclear perimeter compared to the cells of the benign epithelium. This can be explained by the fact that larger nuclei are expected to have longer perimeter. The nuclear perimeter measured by the software is thus expected to be higher in the malignant cells. This finding is similar to imaging analysis of squamous cancer at other sites. [1,13]

The average nuclear circularity observed in malignant cells is higher than in normal cells. This is different from what is noticed visually and what is reported in the literature[2] as malignant cells are known to have irregular borders that can sometimes be angulated. This might be because the algorithm for estimating the circularity is not accurate and may need to be upgraded. Further studies need to be done before a conclusion on the software ability to measure nuclear circularity can be made.

Traditionally, malignant cells including NKSCC cells are known to have hyperchromatic nuclei. This is related to the density of hematoxylin stain in the nucleus. This was clearly revealed by the QuPath software as there was higher nuclear hematoxylin OD (optic density) mean, nuclear hematoxylin OD max (figures 3 and 5) and nuclear hematoxylin OD sum

(table 1) in the malignant cells. The maximum hematoxylin OD is expected to be higher in malignant cells as they tend to have hyperchromatic nuclei. We also noticed higher standard deviation of hematoxylin OD (figure 4). This may be related to the fact that the nucleus of malignant cells may sometimes be vesicular with clumped chromatin.

Max nuclear caliper is different in the cells of the two disease entities. The maximum nuclear caliper is higher in cells of NKSCC than that of benign epithelium. This is expected as the caliper is closely related to the nuclear size.

The area of the malignant cells in tissue sections might be expected to be higher than that of normal cells. This is because nuclear enlargement in malignant cells is often accompanied by increase in cell size (area). We observed this in our results which showed a higher mean cell area of the malignant cells (table 2). In accordance with increase cell area, it is expected that the cell perimeter and maximum cell caliper will also increase. This was seen in the results and in similar works at other sites[1].

We noticed a higher cell eccentricity in NKSCC. This is expected as malignant cells tend to have more irregular structures that may increase its eccentricity.

We only annotated normal areas of benign epithelium in this study. We did not include cornified epithelium commonly noticed in uterine prolapse. Neither did we include areas of parakeratosis that could be seen in some reactive lesions. Generally, keratin is eosinophilic. Consequently, we expect the cytoplasmic eosin OD, likely to be increased in cytoplasm with abundant keratin, to be low in benign epithelium. However, in malignant NKSCC there could be increase in keratin. Even though it is called non-keratinizing, it still contains appreciable amount of keratin but not enough to form keratin pearls. Since normal cells have visually imperceptible keratin, they are expected to have low eosin OD mean and low eosin OD std dev. The malignant cells sometimes have keratin in their cytoplasm and are expected to have a high mean and maximum eosin OD. The eosin OD standard deviation is also expected to be

higher. Our results show increase in mean, maximum and minimum eosin OD in the cytoplasm of the malignant cells. This could be used as a distinguishing factor to identify clusters of NKSCC.

The nucleus cell area ratio of normal epithelium varies from the lower layer to the upper layers. The lower layer has basal cells that have high nucleocytoplasmic ratio while the cells have reduced nuclear size and relatively increased size of the cytoplasm as the cells mature towards the upper layer of the epithelium. Hence the nucleocytoplasmic ratio of the normal epithelium could vary widely. In our study, the full thickness of the epithelium was annotated for analysis to include all the varying nuclei. Squamous cancer cells are generally known to have higher nucleocytoplasmic ratio [1]. This was shown in our work which revealed a higher nucleocytoplasmic ratio for NKSCC cells. However, the significant overlap in the major values of both types of cell in the box plot shown in figure 10 suggests it might be of limited use in distinguishing the cells based on this criteria.

It is essential to emphasize that the other variants of cervical cancer are expected to have differential features unique to them and likely to be different from that of NKSCC. Although small cell carcinoma may be expected to have increased hematoxylin OD and nucleocytoplasmic ratio, small cell carcinoma may have features much different from that of NKSCC as their nuclei is of smaller size and the cytoplasm does not contain much keratin like NKSCC. Clear cell carcinoma may also differ as they have clear cytoplasm and are expected to have much lower cytoplasmic eosin OD mean and eosin OD standard deviation compared to NKSCC.

Keratinizing squamous cell carcinoma will be expected to have features like NKSCC but with much higher cytoplasmic eosin OD mean as it usually has more keratin in its cytoplasm. It will be necessary to study all the variants together to see if it will be possible to objectively distinguish between the entities. Since the degree of keratinization has been linked to disease

behavior and prognosis, the potential of Qu Path software to objectively measure the amount of keratin in the cytoplasm can be explored to authenticate or refute previously held notions about the relationship between keratinization and clinicopathological features of these cancers.

Although numerous models built by machine learning are being developed to make for easier diagnosis of lesions using multiple cellular morphometric features[14], directly utilizing the parameters may be useful to aid tissue diagnosis in special cases.

### **Conclusion**

The QuPath software (version 0.3.2) can be a useful aid to differentiate non-keratinizing squamous cell carcinoma(NKSCC) from normal squamous cervical epithelium especially in fragmented biopsies. NKSCC cells have a higher mean nuclear area, nuclear perimeter, mean nuclear hematoxylin OD, maximum nuclear hematoxylin OD and maximum nuclear caliper. NSKCC cells tend to have higher cell perimeter, maximum cell caliper, cell eccentricity, mean eosin OD, maximum eosin OD and minimum eosin OD than normal cells. Nucleocytoplasmic ratio is higher in the malignant cells but is likely to be of less value in distinguishing it from the malignant cells.

### **Recommendation**

A wider, standardized and more statistically robust study using more advanced data analysis software to handle larger number of cells and to involve other histological types as well as cervical intraepithelial neoplasia is required to assess its clinical use.

### **Availability of data and materials**

The data on all the measurements for each of the 13,774 normal cells and 26,114 cells of non-keratinizing squamous cell carcinoma are available on request.

## References

1. Kumar M, Chatterjee K, Purkait SK, Samaddar D. Computer-assisted morphometric image analysis of cells of normal oral epithelium and oral squamous cell carcinoma. *J Oral MaxillofacPathol JOMFP*. 2017;21(1):24–9.
2. Fischer EG. Nuclear Morphology and the Biology of Cancer Cells. *Acta Cytol*. 2020;64(6):511–9.
3. Prema V, Thomas T, Harikrishnan P, Viswanathan M, Srichinthu KK, Rajkumar K. Morphometric Analysis of Suprabasal Cell Layer in Oral Epithelial Dysplasia: A Computer-assisted Microscopic Study. *J Pharm Bioallied Sci*. 2020 Aug;12(Suppl 1):S204–9.
4. Santoshi CK, Kumar JV, Bhagirath PV, Vinay BH, Prakash YJ. Morphometric analysis of basal cells of oral epithelium in predicting malignant transformation of oral potentially malignant disorders in patients with tobacco chewing habit. *J Oral MaxillofacPathol JOMFP*. 2020;24(3):579.
5. Ananjan C, Jyothi M, Laxmidevi BL, Gopinathan PA, Nazir SH, Pradeep L. Morphometric computer-assisted image analysis of epithelial cells in different grades of oral squamous cell carcinoma. *J Cancer Res Ther*. 2018;14(2):361–7.
6. Prema V, Thomas T, Harikrishnan P, Viswanathan M, Srichinthu KK, Rajkumar K. Morphometric Analysis of Suprabasal Cell Layer in Oral Epithelial Dysplasia: A Computer-assisted Microscopic Study. *J Pharm Bioallied Sci*. 2020 Aug;12(Suppl 1):S204–9.
7. Santoshi CK, Kumar JV, Bhagirath PV, Vinay BH, Prakash YJ. Morphometric analysis of basal cells of oral epithelium in predicting malignant transformation of oral potentially

- malignant disorders in patients with tobacco chewing habit. *J Oral Maxillofac Pathol JOMFP*. 2020;24(3):579.
8. Kesarkar K, Tamgadge A, Peirera T, Tamgadge S, Gotmare S, Kamat P. Evaluation of Mitotic Figures and Cellular and Nuclear Morphometry of Various Histopathological Grades of Oral Squamous Cell Carcinoma: Comparative study using crystal violet and Feulgen stains. *Sultan Qaboos Univ Med J*. 2018 May;18(2):e149–54.
  9. Bankhead P, Loughrey MB, Fernández JA, Dombrowski Y, McArt DG, Dunne PD, et al. QuPath: Open source software for digital pathology image analysis. *Sci Rep*. 2017 Dec 4;7(1):16878.
  10. Allemani C, Matsuda T, Di Carlo V, Harewood R, Matz M, Nikšić M, et al. Global surveillance of trends in cancer survival 2000-14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *Lancet Lond Engl*. 2018 Mar 17;391(10125):1023–75.
  11. Arbyn M, Weiderpass E, Bruni L, Sanjosé S de, Saraiya M, Ferlay J, et al. Estimates of incidence and mortality of cervical cancer in 2018: a worldwide analysis. *Lancet Glob Health*. 2020 Feb 1;8(2):e191–203.
  12. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018 Nov;68(6):394–424.
  13. Lu C, Lewis JS, Dupont WD, Plummer WD, Janowczyk A, Madabhushi A. An oral cavity squamous cell carcinoma quantitative histomorphometric-based image classifier of nuclear morphology can risk stratify patients for disease-specific survival. *Mod Pathol Off J U S Can Acad Pathol Inc*. 2017 Dec;30(12):1655–65.

14. Herrington CS, Poulson R, Pillay N, Bankhead P, Coates PJ. Recent Advances in Pathology: the 2022 Annual Review Issue of The Journal of Pathology. *J Pathol.* 2022;257(4):379–82.

UNDER PEER REVIEW