International Journal of Applied Science
ISSN (Online): 2208-2182
Volume 11 Issue 04 December 2025
DOI: https://doi.org/10.69980/54m2ph07

MORPHOLOGICAL PREDICTORS OF BREAST TUMOUR MALIGNANCY BASED ON CELLULAR NUCLEAR CHARACTERISTICS

Dr. Sarah L. Mitchell[1*], Dr. Ahmed R. El-Sayed[2]

[1]Department of Pathology and Laboratory Medicine, University of Toronto, Toronto, Ontario, Canada, Email: sarah.mitchell@utoronto.ca*
[2]Department of Biomedical Engineering, Cairo University, Giza, Egypt
Email: ahmed.elsayed@cu.edu.eg

Abstract*—Breast cancer (BC) remains one of the leading causes of cancer-related morbidity and mortality among women worldwide. Quantitative assessment of cellular nuclear morphology has emerged as a promising approach for improving the accuracy and objectivity of breast tumour diagnosis while supporting computational pathology and artificial intelligence-assisted diagnostic systems. A secondary analytical cross-sectional study was conducted using the publicly available BC Wisconsin (Diagnostic) Dataset comprising 569 breast tumour cases, including 357 benign and 212 malignant lesions. Thirty nuclear morphological characteristics were evaluated and categorized into mean, standard error, and worst measurements. Descriptive statistics, Welch's independent samples t-test with Benjamini–Hochberg correction, Pearson correlation analysis, and multivariable binary logistic regression were performed to identify independent predictors of breast tumour malignancy. Malignant tumours demonstrated significantly greater nuclear size, irregularity, and structural complexity than benign tumours across most morphological variables. Worst nuclear characteristics exhibited the strongest discriminatory ability, while correlation analysis revealed strong associations among several nuclear size and boundary-related features. After adjustment for multicollinearity, area worst, texture worst, perimeter standard error, smoothness worst, concavity worst, concave points standard error, symmetry worst, and compactness standard error were identified as independent predictors of breast tumour malignancy. Quantitative nuclear morphological characteristics provide reliable and objective indicators for distinguishing benign from malignant breast tumours. The identified independent predictors may support computer-assisted diagnosis, digital pathology, and artificial intelligence-based clinical decision-support systems, contributing to improved diagnostic consistency and advancing precision oncology.*

Keywords: *Breast cancer; Nuclear morphology; Histopathology; Logistic regression; Digital pathology*

1. INTRODUCTION

Breast cancer (BC) is still one of the most common types of cancer among women globally and a significant public health problem because of its high incidence and mortality rate. Diagnosis early and correct plays a pivotal role in better treatment choice, less disease progression and better survival of the patient. Histopathology is still the mainstay for the diagnosis of breast tumours as it allows direct evaluation of cellular and tissue architecture. Pathologists consider nuclear morphology as one of the most reliable indicators of tumors' behavior, because along with the process of malignant transformation, the size, shape, organization of chromatin and irregularity of nucleus structure also change (Katayama et al., 2022). The molecular mechanisms responsible for the morphological alterations are closely related to the progression of BC and can serve as valuable diagnostic and prognostic markers (Heng et al., 2017). The quantitative evaluation of the nuclear features obtained from the FNA and H&E images have also shown that there is an ability to distinguish between benign breast lesions and malignant ones with a better level of objectivity (Rajbongshi et al., 2018). In recent years, the development of digital image analysis has allowed us to accurately measure the morphometric features of the cell nucleus, which help us better predict the aggressiveness of the tumour and clinical outcomes (Whitney et al., 2018). Likewise, nuclear shape changes have been studied in the context of cancer and shown to be a fundamental hallmark of malignant cells (Fischer, 2020). Correlation of nuclear morphometric patterns has also been established with tumour biomarkers and hence supports the importance of quantitative nuclear assessment in the diagnosis of BC (Rawat et al., 2018). In addition, the nuclear shape and orientation properties have been shown to be of prognostic significance, and to predict survival in early-stage BC patients (Lu et al., 2018). In recent times, deep learning-based methods have been developed to allow for completely automated nuclear pleomorphism assessment, demonstrating the increasing role of computational pathology in clinical BC diagnosis (Mercan et al., 2022).

Although there has been significant advancement in the field of digital pathology, subjective visual interpretation is still partially used in conventional histopathological interpretation, leading to inter-observer variability in nuclear atypia evaluation and tumour grading. While computer-assisted image analysis has made quantitative analysis of nuclear morphology easier, many of the approaches have focused on the automated processing of images and have not been developed systematically to determine the relative contribution of individual nuclear characteristics to the malignancy of breast tumours (Das et al., 2020). Furthermore, the disease BC is very heterogeneous in terms of its pathological subtypes, and thus identifying universally valid morphological predictors is challenging (Cserni et al., 2021). Biological mechanisms underlying nuclear morphological abnormalities are also not completely understood, and therefore the translation of nuclear morphological observations to clinically interpretable diagnostic markers is still limited (Singh & Lele, 2022). The traditional morphological classification systems often combine several pathological features and do not fully consider the independent predictive value of each quantitative nuclear feature (Masood, 2016). Additionally, there are few secondary data analyses that have sought to explore reproducible nuclear predictors that can be applied to computational pathology and machine learning using standardized publicly available datasets (Turkki et al., 2019).

Standardized open BC datasets offer a chance to objectively analyze nuclear morphology with reproducible analytical methods and are less subject to observer-dependent variation. Quantitative analysis of nuclear morphology (QANM) has shown significant promise in the prediction of therapeutic response and precision oncology with artificial intelligence (AI) based approaches (Dodington et al., 2021). Combining morphological data with explainable machine learning approaches might increase the transparency of diagnostics and could help develop clinically interpretable prediction models (Binder et al., 2021). Furthermore, cell morphology has been identified as a strong predictor of the state of cancer cells, meaning that quantitative measurements of nuclear morphology can be used to gain more biologically relevant information than current visual assessments (Alizadeh et al., 2020). Given the ongoing relevance of morphological classification in combination with molecular profiling, reproducible nuclear predictors might be useful in more objective diagnostic decision-making and could be used to develop AI-assisted digital pathology workflows in the future for the management of BC (do Nascimento & Otoni, 2020).

1.1 Research Objectives

1.To evaluate the association between cellular nuclear morphological characteristics and breast tumour malignancy.

2.To identify the most significant nuclear predictors distinguishing benign from malignant breast tumours.

3.To assess the predictive performance of selected nuclear characteristics using statistical classification models.

2. METHODOLOGY

2.1 Study Design

The design of this study was secondary analytical cross sectional with the aim of studying the relationship between cell nuclear morphological features and malignancy in breast tumour. Secondary data analysis was deemed suitable as the study was not aimed at recruiting participants nor intervening in clinical care. Using the analytical approach, the nuclear morphological characteristics of both benign and malignant breast tumours were compared and independent nuclear predictors of malignancy were identified by statistical modelling.

2.2 Data Source

The data used in this study were obtained from a publicly available diabetes prediction source (UCI Machine Learning, 2016). The dataset is quantitative measurements obtained from digitized images of fine-needle aspiration (FNA) samples taken from breast masses. Each record corresponds to a patient with a known benign or malignant diagnosis, along with thirty quantitative nuclear morphological variables that characterize the nuclear size, shape, texture and boundary characteristics. The dataset has been extensively utilized in the field of BC diagnosis research, due to the high quality of the morphological measurements as well as the well-defined diagnostic classification.

2.3 Study Population and Sample Size

The final analytical data set consisted of 569 breast tumour cases. Every observation was from a different patient whose diagnostic data were complete and whose nuclear morphology data were obtained. To ensure analytical consistency, records with incomplete diagnostic information or records lacking values for some variables were removed during the preparation of the data sets. The data was publicly available with the entire eligible data included in the final analysis without any sampling.

2.4 Study Variables

The outcome variable was the diagnosis of a breast tumour; either benign or malignant. Thirty quantitative nuclear morphological features (independent variables) were obtained by digitizing the microscopic images of the fine-needle aspiration specimens. These variables have been used to describe various parameters of nuclear morphology, such as size, shape, surface texture, boundary irregularity, symmetry and structural complexity and were analyzed as mean measurements, standard error measurements and worst measurements for comparative and predictive analysis.

2.5 Nuclear Morphological Variables

BC Wisconsin (Diagnostic) Dataset is a thirty quantitative features dataset describing the nuclei morphology of breast tissue specimens from digitized FNA images. These measurements describe various aspects of nuclear morphology such as nuclear size, shape, texture, smoothness, compactness, concavity, symmetry and structural complexity. For systematic assessment, the parameters were divided into three groups based on the measurement technique: mean nuclear characteristics, standard error nuclear characteristics and worst nuclear characteristics.

2.5.1 Mean Nuclear Characteristics

The mean nuclear characteristics are obtained by averaging all the measurements from all the nuclei found in each specimen and indicate the general morphology of the tumour cells. The variables evaluated were radius mean, texture mean, perimeter mean, area mean, smoothness mean, compactness mean, concavity mean, concave points mean, symmetry mean and fractal dimension mean.

2.5.2 Standard Error Nuclear Characteristics

The standard error nuclear characteristics illustrate the variations of nuclear measurements within each specimen and give an idea of morphological heterogeneity of the tumour cells. Evaluated variables were: radius SE, texture SE, perimeter SE, area SE, smoothness SE, compactness SE, concavity SE, concave points SE, symmetry SE and fractal dimension SE.

2.5.3 Worst Nuclear Characteristics

Worst nuclear characteristics are the most extreme or abnormal measurement for each nuclear feature, and are representative of the most extreme nuclear atypia within the specimen. The variables of interest which were evaluated were radius worst, texture worst, perimeter worst, area worst, smoothness worst, compactness worst, concavity worst, concave points worst, symmetry worst, and fractal dimension worst.

2.6 Statistical Analysis

Data were entered into MS Excel. Descriptive statistics were used to summarize the distribution of breast tumour diagnosis, the continuous variables were presented as mean ± standard deviation, while the

categorical variables were summarized as frequencies and percentages. Welch's independent samples t-test was used to compare benign and malignant tumours and the Benjamini–Hochberg false discovery rate adjustment was used to correct for multiple comparisons. Cohen's d was used to determine the effect size of group differences. Pearson correlation analysis was used to examine the relationship between nuclear morphological traits and multicollinearity was checked by variance inflation factor 'VIF'. Multivariable binary logistic regression with backward variable selection was used to identify independent nuclear predictors of breast tumour malignancy, which were reported as adjusted odds ratios (ORs) and associated 95% confidence intervals (CIs). Statistically significant test results were set at a two-tailed P-value < 0.05.

2.7 Ethical Considerations

The secondary data that was used in this study was publicly available and was completely de-identified. The analysis did not require formal ethical approval or informed consent as there was no direct recruitment of human participants or clinical intervention, nor was there access to confidential patient records. The data set was solely used for academic research in compliance with the public availability of the data set and relevant policies for data use.

3. RESULTS

3.1 Distribution of Breast Tumour Cases

The final analysis on the distribution of benign and malignant breast lesions included 569 breast tumour cases. It is important to know the diagnostic makeup of the study population, as this gives context to the subsequent comparative analysis of the nuclear morphological characteristics, and ensures an adequate representation of both diagnostic categories for statistical analysis. Figure 1 shows the distribution of the breast tumour cases by diagnosis.

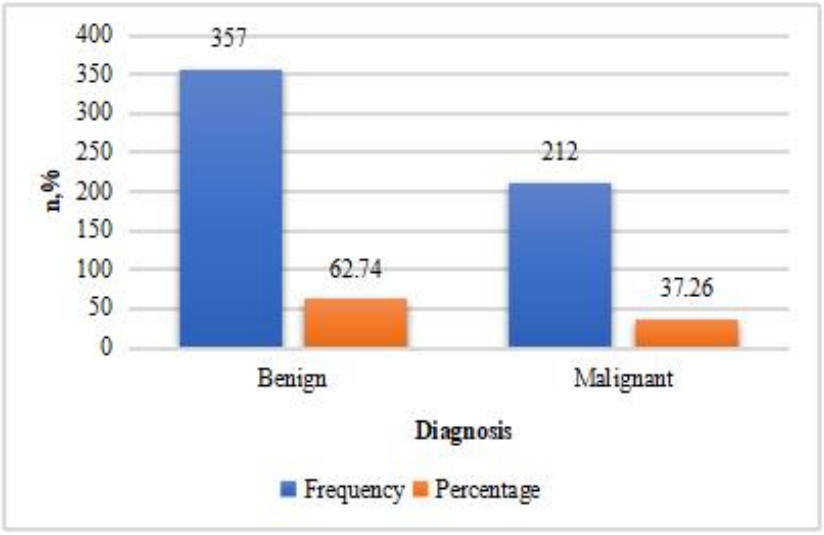

Figure 1. Distribution of breast tumour cases according to diagnosis

As shown in Figure 1, there are 357 (62.74%) benign and 212 (37.26%) malignant breast tumours. While the number of benign lesions was higher, there was a significant number of malignant cases as well and meaningful comparisons were made possible between the two diagnostic groups. The results were based on sufficient numbers of observations from both categories and provided a solid basis for assessing differences in nuclear morphological features and the identification of independent predictors for malignancy of breast tumour.

3.2 Mean Nuclear Morphological Characteristics

When the mean nuclear characteristics were compared between benign and malignant breast tumours, significant differences were found in the morphology of the nuclei (Table 2). Compared to benign lesions, malignant tumours showed significantly higher mean values for nuclear size, boundary irregularity and architectural complexity for almost all the variables. Concave points mean (Cohen's d = 2.545), perimeter mean (d = 2.290), radius mean (d = 2.205), area mean (d = 2.076) and concavity mean (d = 2.003) showed the greatest standardized differences. Moderate, although statistically significant differences were also found for compactness, texture, smoothness and symmetry. After multiple comparison adjustment, only fractal dimension mean did not differ significantly between benign and malignant tumours.

Table 2. Comparison of mean nuclear morphological characteristics between benign and malignant breast tumours

Variable	Benign (Mean ± SD)	Malignant (Mean ± SD)	Cohen's d	Adjusted p-value
Radius mean	12.146 ± 1.781	17.463 ± 3.204	2.205	<0.001
Texture mean	17.915 ± 3.995	21.605 ± 3.780	0.942	<0.001
Perimeter mean	78.075 ± 11.807	115.365 ± 21.855	2.290	<0.001
Area mean	462.790 ± 134.287	978.376 ± 367.938	2.076	<0.001
Smoothness mean	0.0925 ± 0.0134	0.1029 ± 0.0126	0.793	<0.001
Compactness mean	0.0801 ± 0.0338	0.1452 ± 0.0540	1.535	<0.001
Concavity mean	0.0461 ± 0.0434	0.1608 ± 0.0750	2.003	<0.001
Concave points mean	0.0257 ± 0.0159	0.0880 ± 0.0344	2.545	<0.001
Symmetry mean	0.1742 ± 0.0248	0.1929 ± 0.0276	0.723	<0.001
Fractal dimension mean	0.0629 ± 0.0067	0.0627 ± 0.0076	−0.027	0.821

3.3 Standard Error Nuclear Morphological Characteristics

Table 3 shows that there was statistically considerable more variability in several nuclear characteristics among malignant than benign tumours. The biggest differences were between radius standard error (Cohen's d = 1.422), perimeter standard error (d = 1.382) and area standard error (d = 1.353), all indicating the greater heterogeneity within malignant cell nuclei. The texture, smoothness, symmetry, and fractal dimension standard errors showed relatively small or statistically non-significant differences, while the compactness, concavity, and concave points standard errors revealed significant differences.

Table 3. Comparison of standard error nuclear morphological characteristics between benign and malignant breast tumours

Variable	Benign (Mean ± SD)	Malignant (Mean ± SD)	Cohen's d	Adjusted p-value
Radius SE	0.284 ± 0.113	0.609 ± 0.345	1.422	<0.001
Texture SE	1.220 ± 0.589	1.211 ± 0.483	−0.017	0.864
Perimeter SE	2.000 ± 0.771	4.324 ± 2.569	1.382	<0.001
Area SE	21.135 ± 8.844	72.672 ± 61.355	1.353	<0.001
Smoothness SE	0.0072 ± 0.0031	0.0068 ± 0.0029	−0.139	0.117
Compactness SE	0.0214 ± 0.0164	0.0323 ± 0.0184	0.633	<0.001
Concavity SE	0.0260 ± 0.0329	0.0418 ± 0.0216	0.542	<0.001
Concave points SE	0.0099 ± 0.0057	0.0151 ± 0.0055	0.923	<0.001
Symmetry SE	0.0206 ± 0.0070	0.0205 ± 0.0101	−0.013	0.887
Fractal dimension SE	0.0036 ± 0.0029	0.0041 ± 0.0020	0.161	0.049

3.4 Worst Nuclear Morphological Characteristics

The worst nuclear parameters had the best discriminatory power between benign and malignant breast tumours (Table 4). The values of all the worst nuclear measurements were significantly higher in the malignant tumours, especially the nuclear size and contour irregularity. The largest differences were found for the concave points worst (Cohen's d = 2.693), perimeter worst (d = 2.598), radius worst (d = 2.544), and area worst (d = 2.230). The texture, smoothness, compactness, concavity, symmetry and fractal dimension were also shown to have significant increase which underscored the role of the extreme abnormal nuclei in malignant breast lesions.

Table 4. Comparison of worst nuclear morphological characteristics between benign and malignant breast tumours

Variable	Benign (Mean ± SD)	Malignant (Mean ± SD)	Cohen's d	Adjusted p-value
Radius worst	13.380 ± 1.981	21.135 ± 4.284	2.544	<0.001
Texture worst	23.515 ± 5.494	29.318 ± 5.435	1.061	<0.001
Perimeter worst	87.006 ± 13.527	141.370 ± 29.457	2.598	<0.001
Area worst	558.899 ± 163.601	1422.286 ± 597.968	2.230	<0.001
Smoothness worst	0.1250 ± 0.0200	0.1448 ± 0.0219	0.960	<0.001
Compactness worst	0.1827 ± 0.0922	0.3748 ± 0.1704	1.513	<0.001

Concavity worst	0.1662 ± 0.1404	0.4506 ± 0.1815	1.812	<0.001
Concave points worst	0.0744 ± 0.0358	0.1822 ± 0.0463	2.693	<0.001
Symmetry worst	0.2702 ± 0.0417	0.3235 ± 0.0747	0.945	<0.001
Fractal dimension worst	0.0794 ± 0.0138	0.0915 ± 0.0216	0.707	<0.001

To illustrate the relative size of the differences between the groups, the standardized effect sizes of the poorest nuclear morphological characteristics are shown in Figure 2.

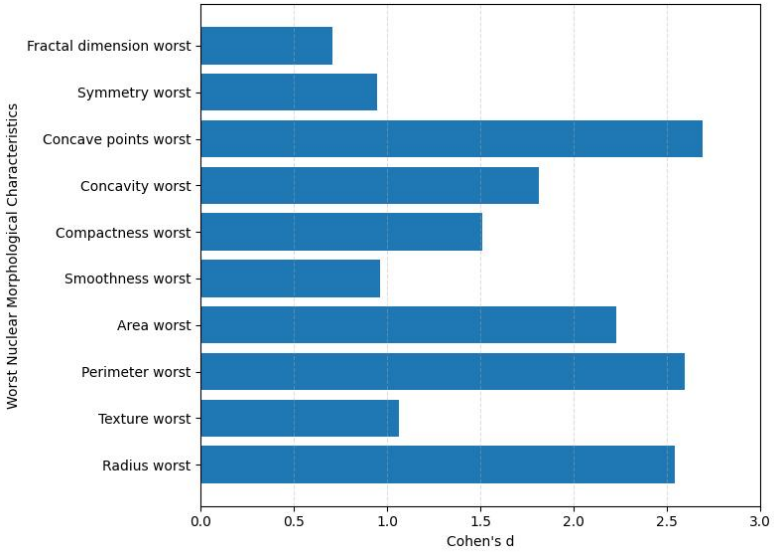

Figure 2. Effect Size of Worst Nuclear Morphological Characteristics

The graphical representation is made to complement the statistical results, to show the relative importance of each worst nuclear characteristic to the discrimination between benign and malignant breast tumours so that they can be compared more easily with respect to their discriminatory strengths for the evaluated variables.

3.5 Correlation Analysis of Nuclear Morphological Characteristics

The multiple nuclear morphological variables showed strong positive correlations as revealed by Pearson correlation analysis (Table 5). The most correlated parameters among all parameters were radius mean and perimeter mean (r = 0.998), and radius worst and perimeter worst (r = 0.994). There were also high positive correlation between radius mean and area mean (r = 0.987), perimeter mean and area mean (r = 0.987), and concavity mean and concave points mean (r = 0.921). These results showed significant associations between the nuclear size measurements and between the nuclear size and the boundary morphology measurements, justifying the use of multivariable models to find independent nuclear size and boundary morphology measurements associated with malignancy.

Table 5. Strongest correlations among nuclear morphological characteristics

Variable 1	Variable 2	Pearson's r
Radius mean	Perimeter mean	0.998
Radius worst	Perimeter worst	0.994
Radius mean	Area mean	0.987
Perimeter mean	Area mean	0.987
Radius worst	Area worst	0.984
Perimeter worst	Area worst	0.978
Radius SE	Perimeter SE	0.973
Radius mean	Radius worst	0.970
Perimeter mean	Perimeter worst	0.970
Concavity mean	Concave points mean	0.921

3.6 Independent Morphological Predictors of Breast Tumour Malignancy

Table 6 displays the eight independent nuclear morphological characteristics that were identified by multivariable logistic regression after adjusting for multicollinearity, which were associated with breast

tumour malignancy. Of these variables, area worst had the highest adjusted OR for association with malignant diagnosis (Adjusted OR = 29,442.94; 95% CI: 798.75–1,085,309.00; p<0.001). There were also significant positive associations with perimeter standard error, texture worst, concave points standard error, concavity worst, smoothness worst and symmetry worst. In contrast, compactness standard error had a highly significant inverse relationship with malignancy (Adjusted OR = 0.10; 95% CI: 0.03 – 0.32; p < 0.001). Multicollinearity of all the retained predictors was acceptable with a range of 1.24 to 4.00 for their VIF values.

Table 6. Independent morphological predictors of breast tumour malignancy identified by multivariable logistic regression

Predictor	Adjusted OR (per SD)	95% CI	p-value	VIF
Area worst	29,442.94	798.75–1,085,309.00	<0.001	3.51
Texture worst	10.11	3.75–27.21	<0.001	1.24
Compactness SE	0.10	0.03–0.32	<0.001	3.36
Smoothness worst	6.87	2.40–19.66	<0.001	1.57
Perimeter SE	26.31	4.18–165.45	<0.001	3.14
Concavity worst	7.99	2.01–31.66	0.003	4.00
Concave points SE	4.39	1.48–12.97	0.008	2.83
Symmetry worst	3.01	1.15–7.88	0.025	1.61

4. DISCUSSION

The aim of the present study was to investigate correlation of the cellular nuclear morphological features with the malignancy level of the breast tumour by quantitative measurement obtained from the BC Wisconsin (Diagnostic) data. Results showed that the mean, standard error, and worst nuclear characteristics of malignant breast tumours were significantly larger, more irregular, and more complex than benign breast tumours in most of the parameters. The worst nuclear features showed the highest discriminatory power among the three types of measurements, suggesting that the most unusual nuclear characteristics provide a very high degree of discriminatory power. Moreover, correlation analysis showed high interrelationships of nuclear size and boundary morphology variables, highlighting the need for multivariable modelling to separate independent predictive variables. Multivariable logistic regression, after multicollinearity adjustment, revealed that area worst, texture worst, perimeter standard error, smoothness worst, concavity worst, concave points standard error, symmetry worst, and compactness standard error were independent predictors of the breast tumour malignancy. From the findings, it can be seen that quantifying nuclear morphology can be used to successfully describe malignant transformation and as an objective indicator to classify breast tumours.

The nuclear morphology differences seen were consistent with the biological evolution of malignant tumours which involve uncontrolled cellular proliferation, resulting in enlarged nuclei, irregular nuclear outlines, increased nuclear chromatin heterogeneity and increased architectural distortion. The higher standard error measurements are also a direct sign of greater cellular heterogeneity which is often related to the aggressiveness of a tumour and its genomic instability. Similarly, the good predictive ability of worst nuclear characteristics indicates that high nuclear abnormalities are more diagnostic than average values, because they reflect the worst nuclear abnormalities in a sample of a tumour. The fact that only a subset of variables remained in the multivariable regression model suggests that many nuclear properties are different for benign vs. malignant lesions, but only specific nuclear morphologic properties contribute to the prediction of malignancy, after accounting for the collinear variation among multiple measurements.

The current results corroborate previous pathological and computational studies that highlighted the role of nuclear morphology in BC diagnosis. Progressive changes in nuclear architecture and cellular morphology are fundamental histopathological characteristics of breast malignant lesions vs non-invasive lesions, reported by Sanati et al. 2019. Likewise, the increased variability of the malignant nuclei could be explained by the single-cell morphological analysis of Zhao et al. (2023), which revealed significant cellular heterogeneity in BC progression. Jia-Mei et al. (2017) also emphasized that quantitative, nuclear morphometric data obtained from histopathological images are reliable diagnostic and prognostic data, which further supports the significant associations found in this study. Similar results have been observed outside BC, with Lu et al. (2017) demonstrating that the quantitative nuclear morphology parameters could be used to stratify oral squamous cell carcinoma patients based on clinical outcomes. This highlights the potential for the wider application of quantitative nuclear morphology parameters in the field of oncology. Similarly, in the field of muscle-invasive bladder cancer, Mi et al. (2021) found that combination of nuclear morphology and tissue architecture created predictive models to identify responses to therapy, showing the strong robustness of morphological biomarkers in different types of tumours. The findings of the present study that these are

independent predictors in BC are further underscored by the importance of artificial intelligence-based digital histopathology in the prognosis and treatment of BC which increasingly focuses on quantitative nuclear morphology (McCaffrey et al. 2024).

The results of this study have clinical and research implications. Quantitative nuclear morphological features have been found to be an objective and reproducible measurement and could be used in conjunction with conventional histopathology in the diagnosis of breast tumours, to minimise observer variance. The independent factors that have been identified can be used in computer-assisted diagnosis systems and AI-assisted pathological workflows to improve the accuracy of tumour classification and aid personalized clinical decision-making. These methods could help streamline diagnostic processes, enable early detection and refine risk assessment in the management of BC.

There are, however, a number of drawbacks. The study used a publicly available secondary data set from a single diagnostic cohort, and the findings might not be applicable to larger patient groups. Furthermore, the analysis was limited to quantitative nuclear morphological variables only and did not include any molecular, genomic, clinical, or treatment-related information which could be used to further enhance the predictive capabilities of the system. Last, due to the cross-sectional design of the data, it was not possible to evaluate long-term clinical outcomes like recurrence or survival.

The nuclear morphological predictors identified in this study should be validated in large, multicentre datasets of different patient cohorts. A combination of quantitative nuclear morphology, molecular biomarkers, radiological features and clinical variables might further improve predictive accuracy, and help create robust multimodal diagnostic models. Moreover, future research with advanced algorithms in artificial intelligence and deep learning could enable automated BC detection and enhance personalized prognostic assessment in clinical practice.

5. CONCLUSION

This study showed, that the quantitative nuclear morphological parameters were able to differentiate between benign and malignant breast tumours by using a standardized secondary data set. There were significant differences between the means of most nuclear measurements, their standard errors and their worst values, thus supporting the hypothesis that malignant tumours have larger, more irregular and morphologically more variable nuclei. The worst nuclear characteristics discriminated best among the evaluated parameters. The independent variables associated with the malignancy of breast tumour were area worst, texture worst, perimeter standard error, smoothness worst, concavity worst, concave points standard error, symmetry worst, and compactness standard error using multivariable logistic regression. These results indicate the clinical importance of objective nuclear morphometric analysis and show that quantitative morphology can be used for proper tumour classification. Moreover, the study highlights the value of publicly available datasets for creating reproducible computational pathology models and the potential for publicly released nuclear morphology data to be adopted into digital pathology and AI-supported decision support systems for more accurate and consistent diagnostics and future precision oncology applications.

REFERENCES

1. Alizadeh, E., Castle, J., Quirk, A., Taylor, C. D., Xu, W., & Prasad, A. (2020). Cellular morphological features are predictive markers of cancer cell state. Computers in biology and medicine, 126, 104044.
2. Binder, A., Bockmayr, M., Hägele, M., Wienert, S., Heim, D., Hellweg, K., ... & Klauschen, F. (2021). Morphological and molecular breast cancer profiling through explainable machine learning. Nature Machine Intelligence, 3(4), 355-366.
3. Cserni, G., Quinn, C. M., Foschini, M. P., Bianchi, S., Callagy, G., Chmielik, E., ... & European Working Group for Breast Screening Pathology. (2021). Triple-negative breast cancer histological subtypes with a favourable prognosis. Cancers, 13(22), 5694.
4. Das, A., Nair, M. S., & Peter, S. D. (2020). Computer-aided histopathological image analysis techniques for automated nuclear atypia scoring of breast cancer: a review. Journal of digital imaging, 33(5), 1091-1121.
5. do Nascimento, R. G., & Otoni, K. M. (2020). Histological and molecular classification of breast cancer: what do we know?. Mastology, 30, 1-8.
6. Dodington, D. W., Lagree, A., Tabbarah, S., Mohebpour, M., Sadeghi-Naini, A., Tran, W. T., & Lu, F. I. (2021). Analysis of tumor nuclear features using artificial intelligence to predict response to neoadjuvant chemotherapy in high-risk breast cancer patients. Breast Cancer Research and Treatment, 186(2), 379-389.
7. Fischer, E. G. (2020). Nuclear morphology and the biology of cancer cells. Acta cytologica, 64(6), 511-519.
8. Heng, Y. J., Lester, S. C., Tse, G. M., Factor, R. E., Allison, K. H., Collins, L. C., ... & Beck, A. H. (2017). The molecular basis of breast cancer pathological phenotypes. The Journal of pathology, 241(3), 375-391.
9. Jia-Mei, C., Li, Y., Xu, J., Gong, L., Lin-Wei, W., Wen-Lou, L., & Liu, J. (2017). Computer-aided prognosis on breast cancer with hematoxylin and eosin histopathology images: A review. Tumor Biology, 39(3).
10. Katayama, A., Toss, M. S., Parkin, M., Sano, T., Oyama, T., Quinn, C. M., ... & Rakha, E. A. (2022). Nuclear

morphology in breast lesions: refining its assessment to improve diagnostic concordance. Histopathology, 80(3), 515-528.
11. Lu, C., Lewis, J. S., Dupont, W. D., Plummer, W. D., Janowczyk, A., & Madabhushi, A. (2017). An oral cavity squamous cell carcinoma quantitative histomorphometric-based image classifier of nuclear morphology can risk stratify patients for disease-specific survival. Modern Pathology, 30(12), 1655-1665.
12. Lu, C., Romo-Bucheli, D., Wang, X., Janowczyk, A., Ganesan, S., Gilmore, H., ... & Madabhushi, A. (2018). Nuclear shape and orientation features from H&E images predict survival in early-stage estrogen receptor-positive breast cancers. Laboratory investigation, 98(11), 1438-1448.
13. Masood, S. (2016). Breast cancer subtypes: morphologic and biologic characterization. Women's Health, 12(1), 103-119.
14. McCaffrey, C., Jahangir, C., Murphy, C., Burke, C., Gallagher, W. M., & Rahman, A. (2024). Artificial intelligence in digital histopathology for predicting patient prognosis and treatment efficacy in breast cancer. Expert review of molecular diagnostics, 24(5), 363-377.
15. Mercan, C., Balkenhol, M., Salgado, R., Sherman, M., Vielh, P., Vreuls, W., ... & Ciompi, F. (2022). Deep learning for fully-automated nuclear pleomorphism scoring in breast cancer. NPJ breast cancer, 8(1), 120.
16. Mi, H., Bivalacqua, T. J., Kates, M., Seiler, R., Black, P. C., Popel, A. S., & Baras, A. S. (2021). Predictive models of response to neoadjuvant chemotherapy in muscle-invasive bladder cancer using nuclear morphology and tissue architecture. Cell Reports Medicine, 2(9).
17. Rajbongshi, N., Bora, K., Nath, D. C., Das, A. K., & Mahanta, L. B. (2018). Analysis of morphological features of benign and malignant breast cell extracted from fnac microscopic image using the pearsonian system of curves. Journal of cytology, 35(2), 99-104.
18. Rawat, R. R., Ruderman, D., Macklin, P., Rimm, D. L., & Agus, D. B. (2018). Correlating nuclear morphometric patterns with estrogen receptor status in breast cancer pathologic specimens. NPJ Breast Cancer, 4(1), 32.
19. Sanati, S. (2019). Morphologic and molecular features of breast ductal carcinoma in situ. The American journal of pathology, 189(5), 946-955.
20. Singh, I., & Lele, T. P. (2022). Nuclear morphological abnormalities in cancer: a search for unifying mechanisms. In Nuclear, chromosomal, and genomic architecture in biology and medicine (pp. 443-467). Cham: Springer International Publishing.
21. Turkki, R., Byckhov, D., Lundin, M., Isola, J., Nordling, S., Kovanen, P. E., ... & Linder, N. (2019). Breast cancer outcome prediction with tumour tissue images and machine learning. Breast cancer research and treatment, 177(1), 41-52.
22. UCI Machine Learning. (2016). Breast Cancer Wisconsin (Diagnostic) Data Set [Data set]. Kaggle. https://www.kaggle.com/datasets/uciml/breast-cancer-wisconsin-data
23. Whitney, J., Corredor, G., Janowczyk, A., Ganesan, S., Doyle, S., Tomaszewski, J., ... & Madabhushi, A. (2018). Quantitative nuclear histomorphometry predicts oncotype DX risk categories for early stage ER+ breast cancer. BMC cancer, 18(1), 610.
24. Zhao, S., Chen, D. P., Fu, T., Yang, J. C., Ma, D., Zhu, X. Z., ... & Shao, Z. M. (2023). Single-cell morphological and topological atlas reveals the ecosystem diversity of human breast cancer. Nature Communications, 14(1), 6796.