HHS Public Access

Author manuscript

Int J Sci Acad Res. Author manuscript; available in PMC 2021 November 24.

Published in final edited form as:

Int J Sci Acad Res. 2021 January ; 2(1): 3081–3086.

Application of Machine Learning Algorithms in Breast Cancer Diagnosis and Classification

Clement G. Yedjou[1], **Solange S. Tchounwou**[2], **Richard A. Aló**[3], **Rashid Elhag**[1], **BereKet Mochona**[4], **Lekan Latinwo**[1]

[1]Department of Biological Sciences, College of Science and Technology, Florida Agricultural and Mechanical University, 1610 S. Martin Luther King Blvd, Tallahassee, FL 32307, United States.

[2]Department of Pathology and Laboratory Medicine. School of Medicine, Tulane University, 1430 Tulane Avenue, New Orleans, LA, 70112, United States.

[3]Department of Computer and Information Science, College of Science and Technology, Florida Agricultural & Mechanical University, 1610 S. Martin Luther King Blvd, Tallahassee, FL 3230, United States.

[4]Department of Chemistry, College of Science and Technology, Florida Agricultural and Mechanical University, 1610 S. Martin Luther King Blvd, Tallahassee, FL 32307, United States.

Abstract

Breast cancer continues to be the most frequent cancer in females, affecting about one in 8 women and causing the highest number of cancer-related deaths in females worldwide despite remarkable progress in early diagnosis, screening, and patient management. All breast lesions are not malignant, and all the benign lesions do not progress to cancer. However, the accuracy of diagnosis can be increased by a combination or preoperative tests such as physical examination, mammography, fine-needle aspiration cytology, and core needle biopsy. Despite some limitations, these procedures are more accurate, reliable, and acceptable, when compared with a single adopted diagnostic procedure. Recent studies have shown that breast cancer can be accurately predicted and diagnosed using machine learning (ML) technology. The objective of this study was to explore the application of ML approaches to classify breast cancer based on feature values generated from a digitized image of a fine-needle aspiration (FNA) of a breast mass. To achieve this objective, we used ML algorithms, collected a scientific dataset of 569 breast cancer patients from Kaggle (https://www.kaggle.com/uciml/breast-cancer-wisconsin-data), analyze and interpreted the data based on ten real-valued features of a breast mass FNA including the radius, texture, perimeter, area, smoothness, compactness, concavity, concave points, symmetry, and fractal dimension. Among the 569 patients tested, 63% were diagnosed with benign breast cancer and 37% were

Corresponding author: Clement G Yedjou, Department of Biological Sciences, College of Science and Technology, Florida A & M University, 1610 S. Martin Luther King Blvd, Tallahassee, FL 32307, United States, Tel: +1-850-599-3908, Fax: +1-850-561-2996, clement.yedjou@famu.edu.

Author Contributions: Conceptualization, C.G.Y., S.S.T., and R.A.A.; Methodology, C.G.Y., R.E., and S.S.T; formal analysis, S.S.T., B.M., C.G.Y and R.E.; investigation, C.G.Y, S.S.T., R.E., R.A.A, L.L., and B.M.; supervision, L.L., and R.A.A; writing-original draft preparation, all authors; writing-reviewing and editing, all authors; All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

diagnosed with malignant breast cancer. Benign tumors grow slowly and do not spread while malignant tumors grow rapidly and spread to other parts of the body.

Keywords

Breast cancer; malignant; benign; machine learning; computer-based learning

1. INTRODUCTION

Breast cancer continues to be the most frequent cancer in females, affecting about one in 8 women and causing the highest number of cancer-related deaths in females worldwide despite remarkable progress in early diagnosis, screening, and patient management[1,2,3]. All breast lesions are not malignant tumors and all the benign lesions do not progress to cancer. The screening and diagnosis of breast cancer have improved by a combination or preoperative tests such as physical examination, mammography, fine-needle aspiration cytology, core needle biopsy, digital breast tomosynthesis, ultrasound, and magnetic resonance[4,5,6,7]. However, these diagnostic procedures have their own limitations. For example, the analysis of mammographic images shows low contract between normal tissues and lesions, which makes it difficult to distinguish malignant masses from benign ones in the images[8,9]. Early detection and accurate prognostication are fundamental to identify patients who could benefit from the treatment and reduce the mortality of cancer diseases[10,11,12].

The use of computer-based learning models has become a predominant area of cancer research. In recent years, several researchers have focused on building systems, both hybrid and fully automatic systems, that could facilitate the diagnosis, prognosis, and prediction of breast cancer outcomes taking a leap using Statistics and Artificial Intelligence. The development of these systems requires different techniques, where the most common are machine learning (ML) algorithms. Several scientific studies have published algorithms and nomograms predicting the pathologic stage of patients with clinically localized cancer or Gleason score upgrading[13,14,15,16,17]. Specifically, ML allows the integration or combination of different layers of data, such as those from medical images, laboratory results, clinical outcomes, biomarkers, and biological features for better prognostication and stratification of patients toward personalized medicine[18,19]. Despite a large scientific interest in this field of research, these prediction models are not frequently used due to limitations in usability and applied computational approaches. Many recent studies have demonstrated that ML approaches have been applied to breast cancer survival prediction, diagnostic ultrasound, and breast cancer outcome prediction with tumor tissue images[20,21,22,23]. Therefore, the objective of this study was to explore the application of ML approaches to classify breast cancer based on feature values generated from a digitized image of a fine-needle aspiration of a breast mass.

2. APPROACHES

2.1. Source of Dataset and Information

We used a publicly available breast cancer dataset from the University of Wisconsin Hospitals, Madison, Wisconsin, USA. This dataset was generated by Dr. William H. Wolberg (General Surgery Department., University of Wisconsin, Clinical Sciences Center, Madison, WI 53792)[24], and consisted of 569 breast cancer patients available on UCI Machine Learning Repository: https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+%28Diagnostic%29.

2.2. Machine Learning Methods

This study was based on Machine learning (ML) algorithms to analyze the dataset of 569 patients with breast cancer and thereby interpreting results. ML is a branch of artificial intelligence (AI) that is used to classify data based on models which have been developed and for predictive analytics, in particular breast cancer[25,26]. It provides tools by which large quantities of data can be automatically analyzed. In the case of the present study, we utilized ML algorithms and collected a scientific dataset of breast cancer patients from Kaggle (https://www.kaggle.com/uciml/breast-cancer-wisconsin-data) and interpreted these data based on different features. The features were computed from a digitized image of a fine needle aspirate (FNA) of a breast mass. Ten (10) real-valued features including: [1] radius (*mean of distances from center to points on the perimeter*), [2] texture (*standard deviation of gray-scale values*), [3] perimeter, [4] area, [5] smoothness (*local variation in radius lengths*), [6] compactness (*perimeter^2 / area - 1.0*), [7] concavity (*severity of concave portions of the contour*), [8] concave points (*number of concave portions of the contour*), [9] symmetry, and [10] fractal dimension *(“coastline approximation” - 1*) were computed for each cell nucleus. Specifically, we used the features to differentiate between benign and malignant tumors.

RESULTS

The data presented in this manuscript are available on UCI Machine Learning Repository: https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+%28Diagnostic%29. In this study, 569 patients with breast cancer were diagnosed at the Wisconsin Hospital. Among the 569 patients diagnosed with breast cancer, 63% were benign and 37% were malignant (Fig. 1).

The geometrical and textural features of the most precise core of biopsy were considered and computed in this study (Table 1). As seen in table 1, the geometrical features and textural features are accurate analyses obtained from a digitized image of a fine needle aspirate (FNA) of a breast mass. These features represent simplest attributes of breast cancer images, and they are important for breast cancer analysis. The mean value of each feature for benign tumor (non-cancerous) is lower when compared to each feature for malignant tumor (cancerous), suggesting that malignant tumor spread to the other parts of the body. Taken together, features as seen in Table 1 below allow us to differentiate between benign and malignant.

Seven (7) real-valued features including radius, texture, perimeter, area, compactness, concavity, and concave points of the cell image allow us to differentiate between benign and malignant. Out of the 7 real-valued features, we selected three (3) features (area, perimeter, and radius) and constructed bar graphs to illustrate that breast cancer can be classified based on the values generated (Fig. 2, 3, 4).

Fig. 2 shows the number of patients with benign tumor and/or malignant tumor relatively to the area of the cell image. The lower the value of the area of the cell image indicates benign breast cancer; suggestive the tumor did not spread to other parts of the human body. The higher value of the area of the cell image indicates that the breast cancer has spread to other parts of the human body.

Fig. 3 shows the number of patients with benign tumors and/or malignant tumors relatively to the perimeter of the cell image. The lower value of the perimeter of the cell image indicates benign breast cancer. The higher value of the perimeter of the cell image indicates malignant breast cancer, suggestive that breast cancer has spread to other parts of the human body.

Fig. 4 shows the number of patients with benign tumors and/or malignant tumors relatively to the radius of the cell image. The lower the value of the radius of the cell image indicates benign breast cancer; suggestive the tumor did not spread to other parts of the human body. The higher value of the radius of the cell image indicates that the breast cancer has spread to other parts of the human body.

Three real-valued features including smoothness, symmetry, and fractual dimension of the cell image do not indicate a particular preference of one diagnosis over the other. Out of the 3 features, we selected 2 features (symmetry and fractual dimension) and constructed bar graphs to illustrate that breast cancer cannot be classified based on the values generated (Fig. 5 and 6).

Fig. 5 shows the number of patients with benign tumor and/or malignant tumor relatively to the symmetry of the cell image.

Fig. 6 shows the number of patients with benign tumor and/or malignant tumor relatively to the fractual dimension of the cell image.

Overall, smaller mean values of radius, texture, perimeter, area, compactness, concavity, and concave points of the cell image tend to indicate benign tumors as seen in table 1. Larger mean values of radius, perimeter, area, compactness, concavity, and concave points of the cell image tend to indicate malignant tumors. Mean values of smoothness, symmetry, and fractual dimension of the cell image do not indicate a particular preference of one diagnosis over the other as seen in Table 1.

Knowing the difference between benign tumors and malignant tumors is very important in the field of medical science and cancer research. In addition, knowing this information may help doctors figure out the best way to manage and treat cancer, in particular breast cancer. Benign tumors grow slowly and do not spread while malignant tumors grow fast and spread

to other parts of the body. Benign tumors are non-cancerous while malignant tumors are cancerous. These tumors can spread to other parts of the body from the point of origin and may destroy adjacent normal cells or tissues. Table 2 below shows the characteristics of normal cells, benign and malignant tumors.

DISCUSSION

Breast cancer is the leading cause of death among middle aged and older women[27]. The present study demonstrates the potential of machine learning (ML) approaches for detecting, analyzing, and classifying breast cancer. Using ML, we were able to evaluate different features of a digitized image of a fine needle aspirate (FNA) of a breast mass made available to researchers by Wolberg et al.[24,28]. The FNA of a breast mass describes the characteristics of the cell nuclei present in the image. FNA is a type of biopsy procedure where a very thin needle is inserted into an area of abnormal tissue or cells with a guide of computerized tomography (CT) scan or ultrasound monitors[29,30,31]. The collected sample is then transferred to a pathologist to study it under a microscope and examine whether cells in the biopsy are normal or abnormal. The results generated based on different feature values indicated that among the 569 patients diagnosed with breast cancer, 63% were benign and 37% were malignant. We found that the mean value of each feature for benign tumor (non-cancerous) is lower when compared to each corresponding feature for malignant tumor (cancerous), suggesting that malignant tumor spread to the other parts of the body (Fig. 2, 3, and 4). Based on these features, we were able to differentiate between benign and malignant tumors (Table 2). Cancer cells have the ability to spread to other parts of the body through the blood and lymphatic systems[27].

Medical researchers and physicians usually identify geometrical features and textural features by viewing biopsy images. Multiple classifiers algorithms are applied on medical datasets to perform predictive analysis about patients and their medical diagnosis[32,33,34,35,36,37]. For example, one analysis using a combination of mammograms and ML approaches has led to an accurate diagnosis of breast cancer[38]. Analyses using histopathological images and automatic grading systems have been applied to successfully determine the Gleason grade of breast cancer, and prostate cancer[39,40]. In addition, several previously published methods have shown the potential of ML methods for automatic breast cancer and prostate cancer detection and grading on digital histopathology images[38,41,42,43,44,45].

CONCLUSION

Breast cancer is one of the leading causes of mortality among women worldwide and it is important to develop novel approaches to screen, diagnose, and treat breast cancer. This paper presents a novel computer-aided diagnosis system for the prediction, diagnosis, and classification of breast cancer using ML. In particular, we discussed the concepts of ML and outlined its application in the classification of breast cancer. Using ML approaches, our findings revealed that among the 569 patients involved in this study, 63% were diagnosed with benign tumors and 37% were diagnosed with malignant tumors. The features including radius, texture, perimeter, area, compactness, concavity, and concave points of the cell

image allow us to differentiate between benign and malignant breast cancer. Other features including smoothness, symmetry, and fractual dimension of the cell image do not indicate a particular preference of one diagnosis over the other. Some benign tumors may progress to malignant tumors. We believe that ML will soon become much more commonplace in many clinical and hospital settings. Our results based on the ML can be translated into tools for future clinical treatment decision-making.

Acknowledgments:

The research described here was made possible with the technical support of Dr. Paul B. Tchounwou at Jackson State University, Jackson, MS, U.S.A. The authors wish to thank the University of Wisconsin Hospitals, Madison for the breast cancer dataset.

Funding:

This work was partly funded by the National Institutes of Health (NCI), Grant # 1U54MD015929-01 at Jackson State University, Jackson, MS, United States and partly by the National Science Foundation, NSF-HRD, Grant # 1201981 through the FGLSAMP at Florida Agricultural and Mechanical University, Tallahassee, FL, United States.

Data Availability Statement:

The breast cancer dataset that support the findings in this paper were made available in Kaggle (https://www.kaggle.com/uciml/breast-cancer-wisconsin-data).

References

1. Nolan J, Dunne SS, Mustafa W, Sivananthan L, Kiely PA, Dunne CP. Proposed hypothesis and rationale for association between mastitis and breast cancer. Med Hypotheses. 2020;144. doi:10.1016/j.mehy.2020.110057
2. DeSantis CE, Ma J, Gaudet MM, et al. Breast cancer statistics, 2019. CA Cancer J Clin. 2019;69(6). doi:10.3322/caac.21583
3. Siegel RL, Miller KD, Jemal A. Cancer statistics of American 2019. CA Cancer J Clin. 2019;69(1).
4. Cai D, Lin T, Jiang K, Sun Z. Diagnostic value of MRI combined with ultrasound for lymph node metastasis in breast cancer: Protocol for a meta-analysis. Med (United States). 2019;98(30). doi:10.1097/MD.0000000000016528
5. Ward RC, Lourenco AP, Mainiero MB. Ultrasound-guided breast cancer cryoablation. Am J Roentgenol. 2019;213(3). doi:10.2214/AJR.19.21329
6. Ortiz-Mendoza CM, Sánchez NAA, Dircio AC. Fine-needle aspiration cytology to identify a rare mimicker of breast cancer: Plasma cell mastitis. Rev Bras Ginecol e Obstet. 2018;40(8). doi:10.1055/s-0038-1666809
7. Rana M, Atri SK, Bhagat S, Rana VM. ROLE OF FINE NEEDLE ASPIRATION CYTOLOGY IN BREAST CANCER SCREENING. J Evol Med Dent Sci. 2019;8(8). doi:10.14260/jemds/2019/115
8. Lewin JM, Hendrick RE, D'Orsi CJ, et al. Comparison of full-field digital mammography with screen-film mammography for cancer detection: Results of 4,945 paired examinations. Radiology. 2001;218(3). doi:10.1148/radiology.218.3.r01mr29873
9. Obenauer S, Luftner-Nagel S, von Heyden D, Munzel U, Baum E, Grabbe E. Screen film vs full-field digital mammography: Image quality, detectability and characterization of lesions. Eur Radiol. 2002;12(7). doi:10.1007/s00330-001-1269-y
10. Fiorica JV Breast Cancer Screening, Mammography, and Other Modalities. Clin Obstet Gynecol. 2016;59(4). doi:10.1097/GRF.0000000000000246
11. Sechopoulos I, Teuwen J, Mann R. Artificial intelligence for breast cancer detection in mammography and digital breast tomosynthesis: State of the art. Semin Cancer Biol. 2021;72. doi:10.1016/j.semcancer.2020.06.002 [PubMed: 32927018]

12. Hamdy FC, Donovan JL, Lane JA, et al. 10-Year Outcomes after Monitoring, Surgery, or Radiotherapy for Localized Prostate Cancer. N Engl J Med. 2016;375(15). doi:10.1056/nejmoa1606220
13. Lughezzani G, Briganti A, Karakiewicz PI, et al. Predictive and prognostic models in radical prostatectomy candidates: A critical analysis of the literature. Eur Urol. 2010;58(5). doi:10.1016/j.eururo.2010.07.034
14. Zhang Q, Zhao K, Song L, et al. A Novel Apoptosis-Related Gene Signature Predicts Biochemical Recurrence of Localized Prostate Cancer After Radical Prostatectomy. Front Genet. 2020;11. doi:10.3389/fgene.2020.586376
15. D'Amico AV, Renshaw AA, Arsenault L, Schultz D, Richie JP. Clinical predictors of upgrading to Gleason grade 4 or 5 disease at radical prostatectomy: Potential implications for patient selection for radiation and androgen suppression therapy. Int J Radiat Oncol Biol Phys. 1999;45(4). doi:10.1016/S0360-3016(99)00260-6
16. Chun FKH, Briganti A, Shariat SF, et al. Significant upgrading affects a third of men diagnosed with prostate cancer: Predictive nomogram and internal validation. BJU Int. 2006;98(2). doi:10.1111/j.1464-410X.2006.06262.x
17. Veltri RW, Miller MC, Partin AW, Poole EC, O'Dowd GJ. Prediction of prostate carcinoma stage by quantitative biopsy pathology. Cancer. 2001;91(12). doi:10.1002/1097-0142(20010615)91:12<2322::AID-CNCR1264>3.0.CO;2-H
18. Zitnik M, Nguyen F, Wang B, Leskovec J, Goldenberg A, Hoffman MM. Machine learning for integrating data in biology and medicine: Principles, practice, and opportunities. Inf Fusion. 2019;50. doi:10.1016/j.inffus.2018.09.012
19. Trivizakis E, Papadakis GZ, Souglakos I, et al. Artificial intelligence radiogenomics for advancing precision and effectiveness in oncologic care (Review). Int J Oncol. 2020;57(1). doi:10.3892/ijo.2020.5063
20. Turkki R, Byckhov D, Lundin M, et al. Breast cancer outcome prediction with tumour tissue images and machine learning. Breast Cancer Res Treat. 2019;177(1). doi:10.1007/s10549-019-05281-1
21. Montazeri M, Montazeri M, Montazeri M, Beigzadeh A. Machine learning models in breast cancer survival prediction. Technol Heal Care. 2016;24(1). doi:10.3233/THC-151071
22. Mihaylov I, Nisheva M, Vassilev D. Application of machine learning models for survival prognosis in breast cancer studies. Inf. 2019;10(3). doi:10.3390/info10030093
23. Wu T, Sultan LR, Tian J, Cary TW, Sehgal CM. Machine learning for diagnostic ultrasound of triple-negative breast cancer. Breast Cancer Res Treat. 2019;173(2). doi:10.1007/s10549-018-4984-7
24. Wolberg WH, Street WN, Mangasarian OL. Machine learning techniques to diagnose breast cancer from image-processed nuclear features of fine needle aspirates. Cancer Lett. 1994;77(2–3). doi:10.1016/0304-3835(94)90099-X
25. Nitta S, Tsutsumi M, Sakka S, et al. Machine learning methods can more efficiently predict prostate cancer compared with prostate-specific antigen density and prostate-specific antigen velocity. Prostate Int. 2019;7(3). doi:10.1016/j.prnil.2019.01.001
26. Dey A Machine Learning Algorithms: A Review. Int J Comput Sci Inf Technol. 2016.
27. Qasem A, Abdullah SNHS, Sahran S, Hussain RI, Ismail F. An accurate rejection model for false positive reduction of mass localisation in mammogram. Pertanika J Sci Technol. 2017;25(S6).
28. Wolberg WH, Street WN, Heisey DM, Mangasarian OL. Computer-derived nuclear features distinguish malignant from benign breast cytology. Hum Pathol. 1995;26(7). doi:10.1016/0046-8177(95)90229-5
29. Tanoue LT. Computed Tomography — An Increasing Source of Radiation Exposure. Yearb Pulm Dis. 2009;2009. doi:10.1016/s8756-3452(08)79173-4
30. Guo R, Lu G, Qin B, Fei B. Ultrasound Imaging Technologies for Breast Cancer Detection and Management: A Review. Ultrasound Med Biol. 2018;44(1). doi:10.1016/j.ultrasmedbio.2017.09.012

31. Okamura Y, Yoshizawa N, Yamaguchi M, Kashiwakura I. Application of Dual-Energy Computed Tomography for Breast Cancer Diagnosis. Int J Med Physics, Clin Eng Radiat Oncol. 2016;05(04). doi:10.4236/ijmpcero.2016.54029
32. Asri H, Mousannif H, Al Moatassime H, Noel T. Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis. In: Procedia Computer Science. Vol 83. ; 2016. doi:10.1016/j.procs.2016.04.224
33. Nicolò C, Périer C, Prague M, et al. Machine Learning and Mechanistic Modeling for Prediction of Metastatic Relapse in Early-Stage Breast Cancer. JCO Clin Cancer Informatics. 2020;(4). doi:10.1200/cci.19.00133
34. Carvalho D, Cruz R. Big data and machine learning in health. Eur J Public Health. 2020;30(Supplement_2). doi:10.1093/eurpub/ckaa040.030
35. Sayeth Saabith AL, Sundararajan E, Abu Bakar A. Comparative Study on Different Classification Techniques for Breast Cancer Dataset. Int J Comput Sci Mob Comput. 2014;3(10).
36. El_Rahman SA. Predicting breast cancer survivability based on machine learning and features selection algorithms: a comparative study. J Ambient Intell Humaniz Comput. 2021;12(8). doi:10.1007/s12652-020-02590-y
37. Ceylan Z Diagnosis of breast cancer using improved machine learning algorithms based on bayesian optimization. Int J Intell Syst Appl Eng. 2020;8(3). doi:10.18201/ijisae.2020363531
38. Akselrod-Ballin A, Chorev M, Shoshan Y, et al. Predicting breast cancer by applying deep learning to linked health records and mammograms. Radiology. 2019;292(2). doi:10.1148/radiol.2019182622
39. He L, Long LR, Antani S, Thoma GR. Histology image analysis for carcinoma detection and grading. Comput Methods Programs Biomed. 2012;107(3). doi:10.1016/j.cmpb.2011.12.007
40. Nir G, Karimi D, Goldenberg SL, et al. Comparison of Artificial Intelligence Techniques to Evaluate Performance of a Classifier for Automatic Grading of Prostate Cancer from Digitized Histopathologic Images. JAMA Netw Open. 2019;2(3). doi:10.1001/jamanetworkopen.2019.0442
41. Haq AU, Li JP, Saboor A, et al. Detection of Breast Cancer through Clinical Data Using Supervised and Unsupervised Feature Selection Techniques. IEEE Access. 2021;9. doi:10.1109/ACCESS.2021.3055806
42. Elsadig MA. A machine learning approach for breast cancer early detection. J Theor Appl Inf Technol. 2021;99(5).
43. Mosquera-Lopez C, Agaian S, Velez-Hoyos A, Thompson I. Computer-Aided Prostate Cancer Diagnosis from Digitized Histopathology: A Review on Texture-Based Systems. IEEE Rev Biomed Eng. 2015;8. doi:10.1109/RBME.2014.2340401
44. Nir G, Hor S, Karimi D, et al. Automatic grading of prostate cancer in digitized histopathology images: Learning from multiple experts. Med Image Anal. 2018;50. doi:10.1016/j.media.2018.09.005
45. Wang D, Foran DJ, Ren J, Zhong H, Kim IY, Qi X. Exploring automatic prostate histopathology image gleason grading via local structure modeling. In: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS. Vol 2015-November.; 2015. doi:10.1109/EMBC.2015.7318936

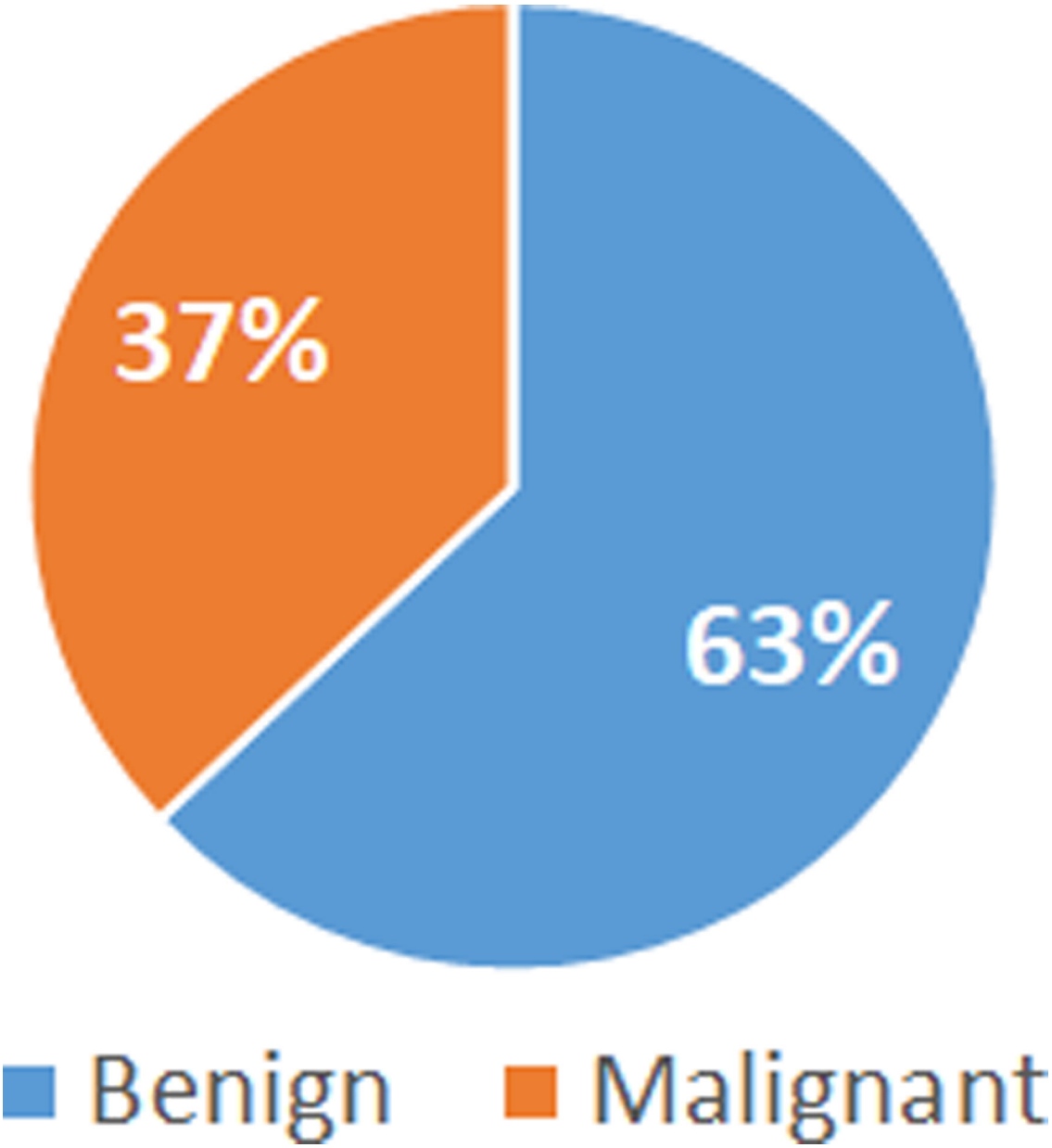


Figure 1:
Percentage of benign and malignant identified among 569 patients with breast cancer

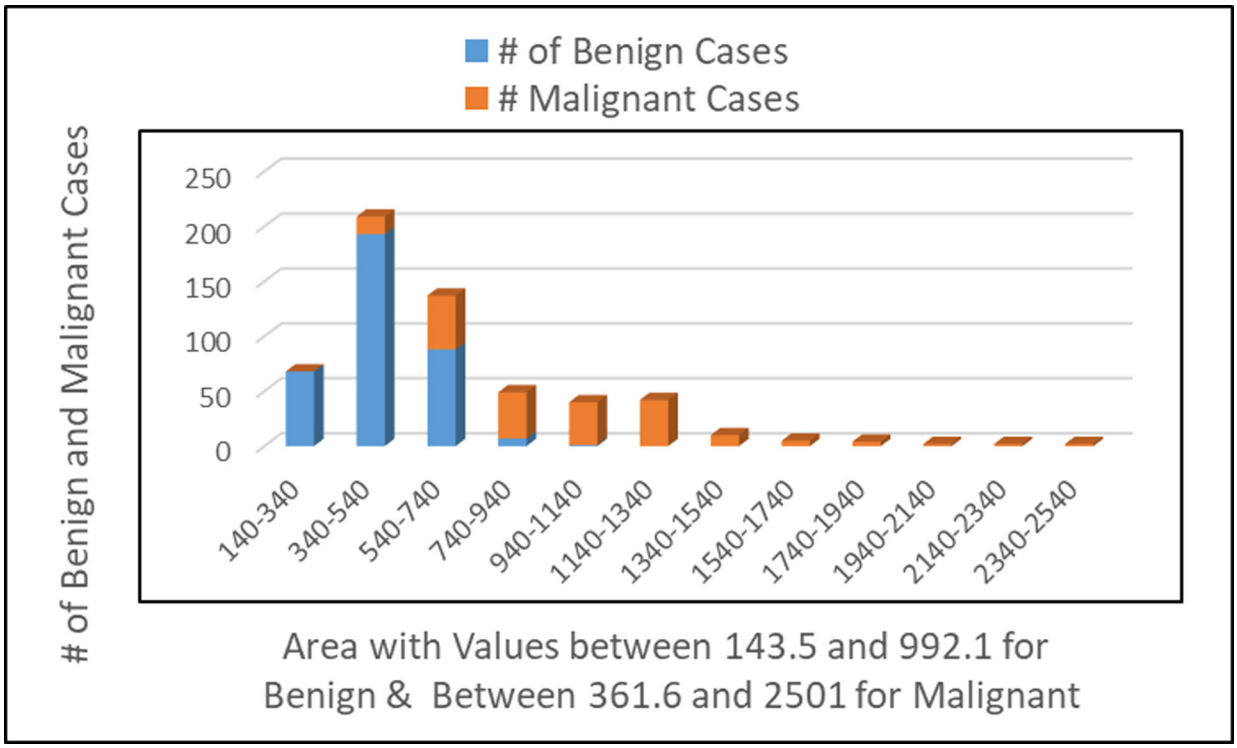


Figure 2:
Number of patients observed with benign and malignant tumors in relationship to the area of the cell image

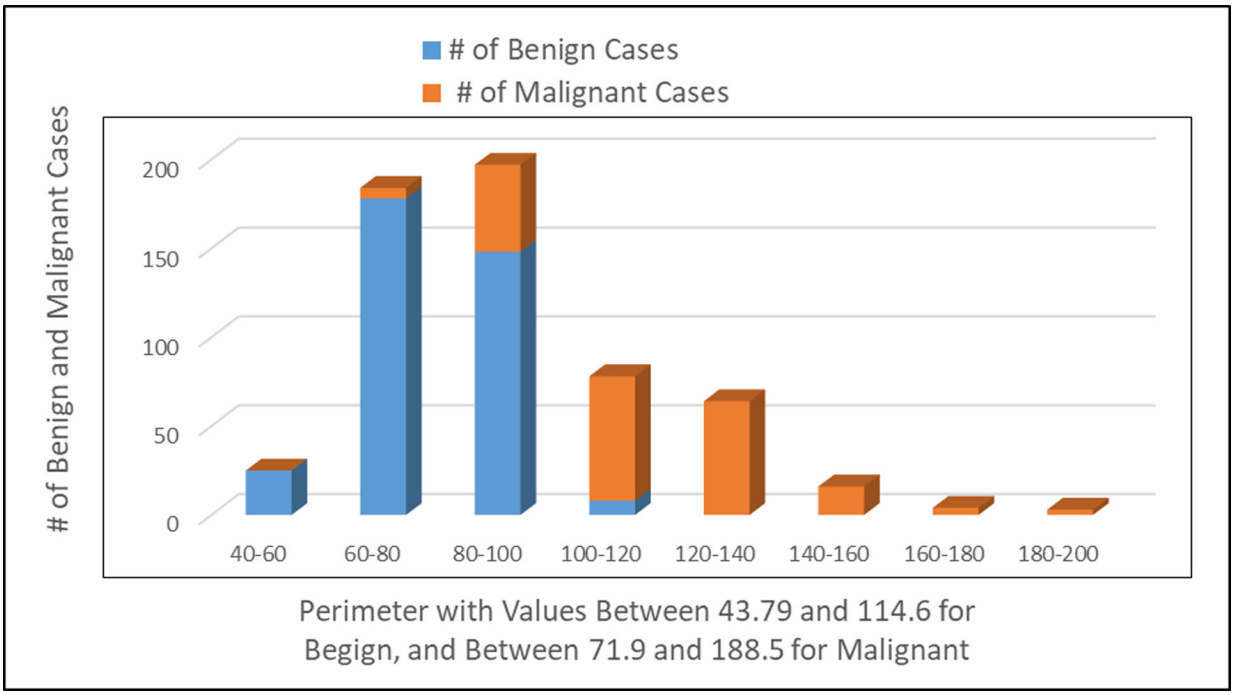


Figure 3:
Number of patients observed with benign and malignant tumors in relationship to the perimeter of the cell image.

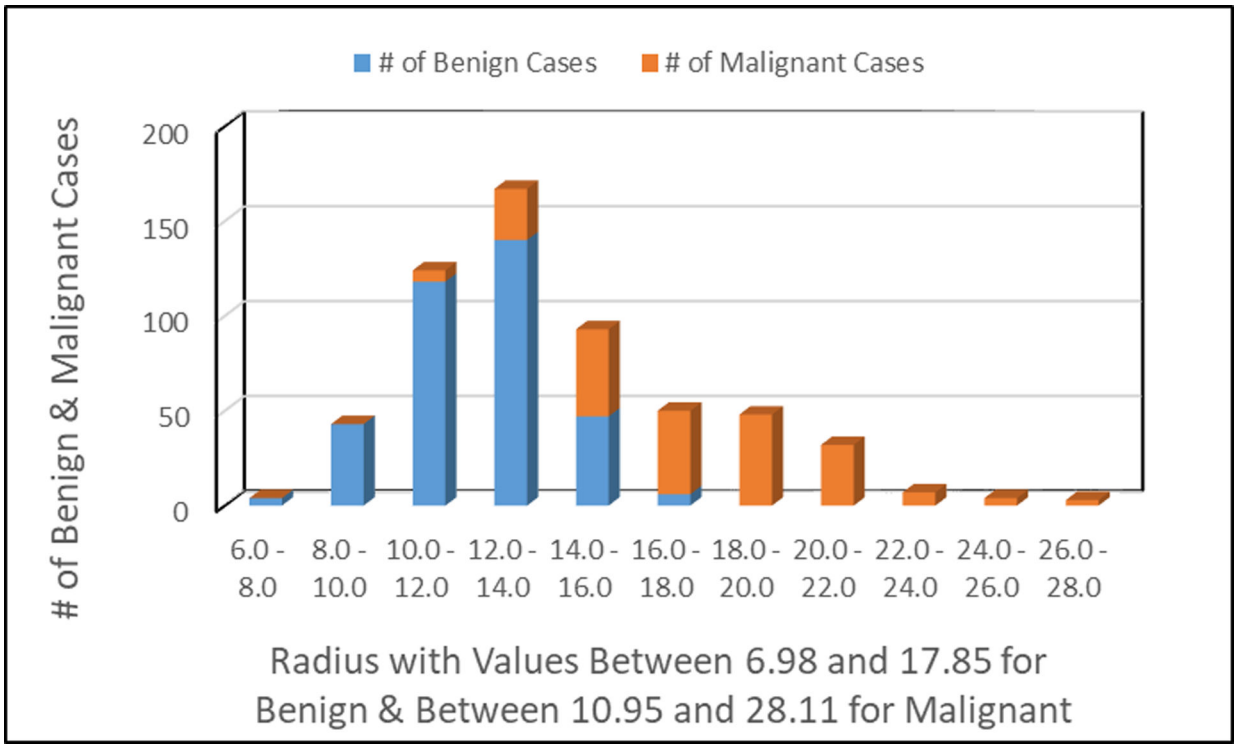


Figure 4:
Number of patients observed with benign and malignant tumors in relationship to the radius of the cell image.

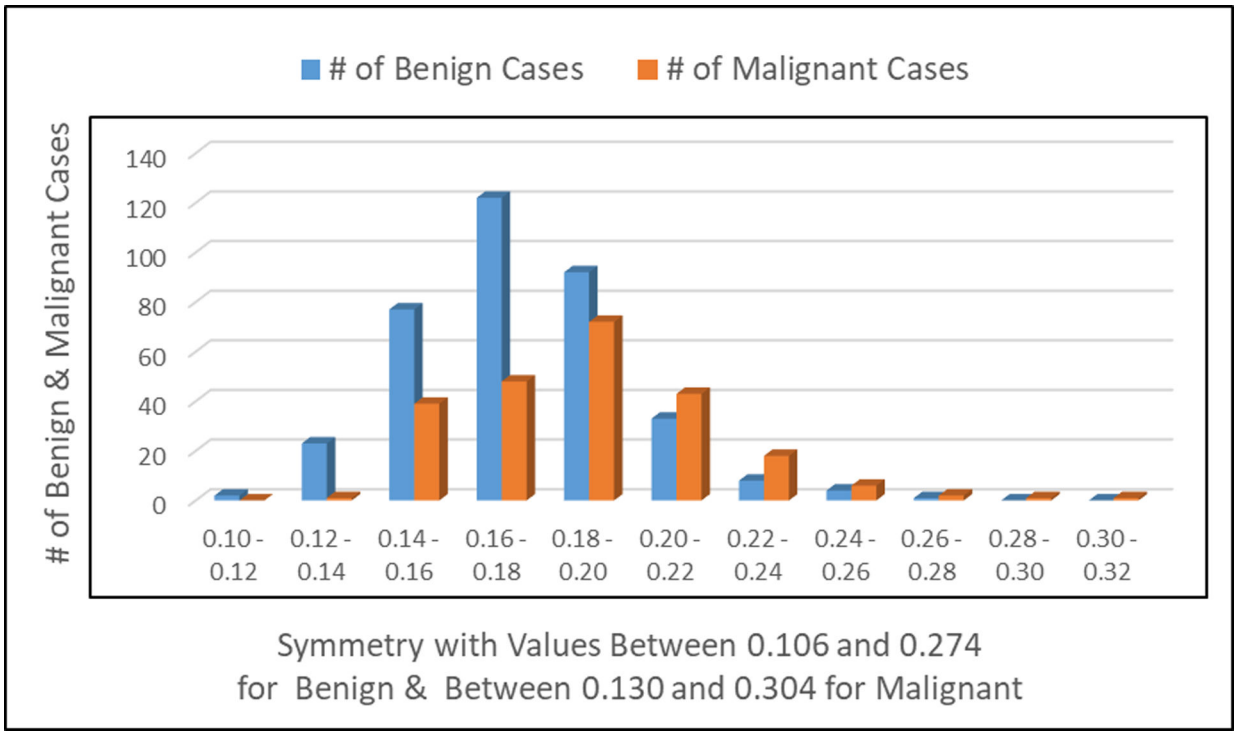


Figure 5:
Number of patients observed with benign and malignant tumors in relationship to the symmetry of the cell image.

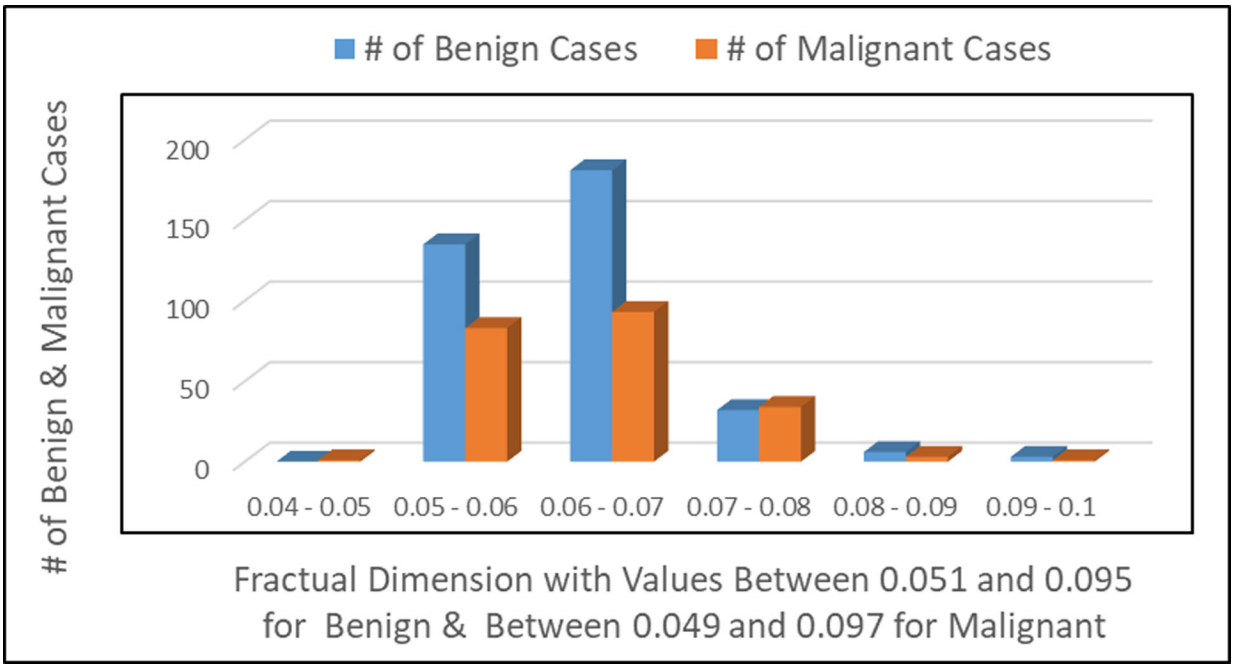


Figure 6:
Number of patients observed with benign and malignant tumors in relationship to the fractual dimension of the cell image.

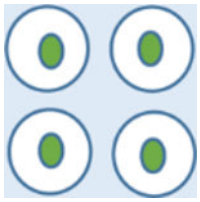
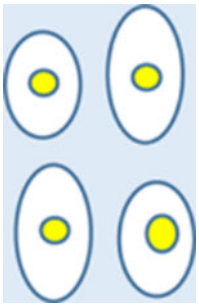
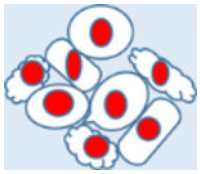
Table 1:

Values of geometrical features and textural features between benign and malignant tumors from a digitized image of a fine needle aspirate (FNA) of a breast mass showing mean values plus and minus standard error values.

FEATURES	BREAST CANCER	
	BENIGN	MALIGNANT
Radius Mean	12.146 ± 0.284	17.463 ± 0.609
Texture Mean	17.914 ± 1.220	21.604 ± 1.210
Perimeter Mean	78.075 ± 2.000	115.365 ± 4.323
Area Mean	462.790 ± 21.135	978.376 ± 72.672
Smoothness	0.092 ± 0.007	0.103 ± 0.006
Compactness	0.080 ± 0.0214	0.145 ± 0.0322
Concavity	0.046 ± 0.0259	0.161 ± 0.0418
Concave Points	0.025 ± 0.009	0.088 ± 0.015
Symmetry	0.174 ± 0.020	0.193 ± 0.020
Fractal Dimension	0.063 ± 0.003	0.063 ± 0.004

Table 2:

Characteristics of normal cells, benign, and malignant tumors

Characteristics	Normal Cells	Benign Tumor cells	Malignant Tumor Cells
Cell Morphology	 Normal cell shape	 Like normal with slight expansion	 Varied in shape and size with large nucleus.
Growth Condition	Grow normally and well-regulated	Grow slowly	Grow rapidly
Spread	Grow in one location	Do not invade surrounding cells, do not invade other parts of the body	Metastasize to other organs through the blood vessels.
Chromosomes	Diploid	Diploid	Aneuploidy
Adherence	Tight	Tight	Loose
Systemic Effects	No	Rare	Yes
Cancer	No	Non-cancerous	Cancerous, spread of tumors to the other parts of the body