BIBLIOGRAPHIC INFORMATION SYSTEM

**Journal Full Title:** Journal of Biomedical Research & Environmental Sciences

**Journal NLM Abbreviation:** J Biomed Res Environ Sci

**Journal Website Link:** https://www.jelsciences.com

**Journal ISSN:** 2766-2276

**Category:** Multidisciplinary

**Subject Areas:** Medicine Group, Biology Group, General, Environmental Sciences

**Topics Summation:** 128

**Issue Regularity:** Monthly

**Review Process type:** Double Blind

**Time to Publication:** 7-14 Days

**Indexing catalog:** Visit here

**Publication fee catalog:** Visit here

**DOI:** 10.37871 (CrossRef)

**Plagiarism detection software:** iThenticate

**Managing entity:** USA

**Language:** English

**Research work collecting capability:** Worldwide

**Organized by:** SciRes Literature LLC

**License:** Open Access by Journal of Biomedical Research & Environmental Sciences is licensed under a Creative Commons Attribution 4.0 International License. Based on a work at SciRes Literature LLC.

Manuscript should be submitted in Word Document (.doc or .docx) through Online Submission form or can be mailed to support@jelsciences.com

**Vision:** Journal of Biomedical Research & Environmental Sciences main aim is to enhance the importance of science and technology to the scientific community and also to provide an equal opportunity to seek and share ideas to all our researchers and scientists without any barriers to develop their career and helping in their development of discovering the world.
Improving Invasive Breast Cancer Care Using Machine Learning Technology

Clement G Yedjou*, Solange S Tchounwou1, Jameka Grigsby3, Keara Johnson4 and Paul B Tchounwou4

1Department of Biological Sciences, College of Science and Technology, Florida Agricultural and Mechanical University, 1610 S. Martin Luther King Blvd, Tallahassee, FL 32307, United States
2Department of Pathology and Laboratory Medicine. School of Medicine, Tulane University, 1430 Tulane Avenue, New Orleans, LA, 70112, United States
3Department of Biological Sciences, School of Arts and Sciences, Alcorn State University, 1000 ASU Drive, Alcorn State, MS 39096, United States
4Department of Biology, College of Science, Engineering and Technology, Jackson State University, 1400 Lynch Street, Box 18750, Jackson, MS 39217, United States

*Corresponding author(s)

Clement G Yedjou, 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
Tel: +1-850-599-3908
Fax: +1-850-561-2996
E-mail: clement.yedjou@famu.edu
DOI: 10.37871/jbres1540
Submitted: 23 August 2022
Accepted: 28 August 2022
Published: 30 August 2022
Copyright: © 2022 Yedjou CG, et al. Distributed under Creative Commons CC-BY 4.0
OPEN ACCESS

ABSTRACT

Breast Cancer (BC) is the most common malignancy in women worldwide. In the United States, the lifetime risk of developing an invasive form of breast cancer is 12.5% among women. BC arises in the lining cells (epithelium) of the ducts or lobules in the glandular tissue of the breast. The goal of the present study was to use Machine Learning (ML) as a novel technology to assess and compare the invasive forms of BC including, infiltrating ductal carcinoma, infiltrating lobular carcinoma, and mucinous carcinoma. To achieve this goal, we used ML algorithms and collected a dataset of 334 BC patients available at https://www.kaggle.com/amandam1/breastcancerdataset and interpreted this dataset based on the form of BC, age, sex, tumor stages, surgery type, and survival rate. Among the 334 patients, 70% were diagnosed with infiltrating ductal carcinoma, 27% with infiltrating lobular carcinoma, and 3% with mucinous carcinoma. Overall, out of 334 BC patients: 64 (19.16%) were in stage I, 189 (56.59%) in stage II, and 81 (24.25%) in stage III. Sixty-six, 67, 96, and 105 patients underwent lumpectomy, simple mastectomy, modified radical mastectomy, and other types of surgery, respectively. The survival rates were 83.4% for stage I, 79.1% for stage II, and 77% for stage III. Findings from the present study demonstrated that ML provides an important tool to curate large amount of BC data, as well as a scientific means to improve BC outcomes.

INTRODUCTION

Breast Cancer (BC) is a group of diseases in which cancer cells are genetically and morphologically different from normal cells and divide in an uncontrolled manner, resulting in a lump or mass [1–3]. Most of breast cancers start in the lobules or in the ducts that connect the lobules to the nipple. Worldwide, BC is the most frequently diagnosed life-threatening cancer and the leading cause of death among women. The development of BC can begin in different areas of the breast, such as the lobules, ducts, or the tissue in between. Adenocarcinoma represents more than 95% of breast cancers [4–6]. The main two histological subtypes of BC are Invasive Lobular Carcinoma (ILC), also called Infiltrating Lobular Carcinoma (ILC), and Invasive Ductal Carcinoma (IDC), also called Infiltrating Ductal Carcinoma (IDC). IDC is the most common type of BC overall, representing approximately 80% of all breast cancers [7,8]. Other pathological subtypes of BC may still appear in patients with IDC. In comparison with IDC, ILC accounts for 10–15% of all BC [9,10], and it is characterized by small, round tumor cells growing in the stroma in a discohesive single-file pattern [11,12]. In addition, ILC is more difficult to detect by standard imaging techniques such as mammography and 18F–FDG–ET [13–15]. ILC is mostly detected in older patients and in advanced stages of BC [16,17]. Furthermore, ILC...
patients display relatively late recurrences and worse long-term survival compared to IDC patients [18–20]. Mucinous BC is a rare type of invasive BC that accounts for less than 2% of all breast cancers. Like other types of invasive BC, mucinous BC begins in the milk duct of the breast before spreading to the tissues around the duct [21]. Most patients with mucinous breast carcinoma present with a palpable breast mass. Mucinous BC grows slowly and may reach a large size at the moment of diagnosis. This can be explained by the fact that the mucinous content of the tumor does not feel firm or solid upon examination [22,23]. It is important to distinguish between the various subtypes of breast cancers because they have different prognoses and treatment implications. The treatment of BC is highly effective when the disease is diagnosed at an early stage. The current treatment options for BC consist of a multidisciplinary approach involving surgical oncology, chemotherapy, radiotherapy, hormone/endocrine therapy, targeted cancer drugs, and bone strengthening drugs (bisphosphonates). Such treatment modalities have been associated with a reduction of BC growth and spread, reduction in breast cancer mortality, and an increase in survival rates, thereby saving the life of BC patients [24]. Researchers are increasingly using Machine Learning (ML) approaches for modelling the progression and treatment of cancer due to its ability to detect key features from complex datasets [25,26]. The application of ML models for the prediction and prognosis of disease development has become an irrevocable part of cancer research. Recent studies in our laboratory have shown that ML is an effective scientific tool to accurately predict, and classify BC patients with benign and malignant tumors [27]. ML holds the potential to transform healthcare and hence opens a world of incredible promise. The goal of the present study was to use ML as a novel technology to assess and compare the invasive forms of BC including, infiltrating ductal carcinoma, infiltrating lobular carcinoma, and mucinous carcinoma.

Materials and Methods

Source of dataset and information

Clinical and pathologic data were extracted from the electronic medical record on BC publicly available on Machine Learning Repository: https://www.kaggle.com/amandam1/breastcancerdataset. The dataset consists of a group of BC patients, who underwent surgical treatment to remove their tumors.

Machine learning methods

This study was based on the application of Machine Learning (ML) algorithms to analyze and interpret a dataset of 334 BC patients. 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 [28,29]. 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 BC patients from Kaggle (https://www.kaggle.com/amandam1/breastcancerdataset) and interpreted this dataset based on different parameters, including patient age at diagnosis, gender, tumor stage, histology, type of surgery, and patient status (alive or dead).

Data selection and collection

We collected a dataset consisting of a group of 334 BC patients, who had surgery to remove their tumors. The dataset consists of the following variables: (1) Age (age at diagnosis); (2) Gender (male/female); (3) Tumor stage (I, II, III); (4) Histology (infiltrating ductal carcinoma, infiltrating lobular carcinoma, and mucinous carcinoma); (5) Surgery type (lumpectomy, simple mastectomy, modified radical mastectomy, other); and (6) Patient status (alive, dead, no reported [no information available whether the patient is alive or dead]).

Results

Invasive breast cancer

Among the 334 BC patients evaluated in this study, 70% of patients were diagnosed with infiltrating ductal carcinoma, also called invasive ductal carcinoma, and 27% of patients were diagnosed with infiltrating lobular carcinoma, also called invasive lobular carcinoma. Also, 3% of patients were diagnosed with mucinous carcinoma. As seen in the case of the present study, invasive ductal carcinoma is the most common type of breast cancer overall, representing over 70% of all breast cancers.

Sex and age at diagnosis

BC is most often found in women, but men also can suffer from it. Among the 334 BC patients evaluated in this study, 330 patients were women aged from 29 to 90 years old, and the remaining four patients were men aged from 44 to 84 years. BC is a rare disease in men. It has been reported that 1% of BC diagnosed in the United States is found in men [30–32]. In the present study, the patients were subsequently divided into thirteen groups of age as follows: 29 to 30 years; 31 to 35 years; 36 to 40 years; 41 to 45 years; 46 to 50 years; 51 to 55 years; 56 to 60 years; 61 to 65 years; 66 to 70 years; 71 to 75 years; 76 to 80 years; 81 to 85 years; and 86 to 90 years. As seen in figure 1, the prevalence of BC increases with age. Also, a very large proportion of BC patients were diagnosed between the age of 46 and 65 compared to the age ranges of 29–45 and 66–90. Although BC can strike at any age, younger females in general are not at high risk of developing BC.

Stages of breast cancer in relation to the age

Staging describes the spread of cancer at the time of diagnosis, and is used to determine the treatment options. Here, 64 out of 334 BC patients were diagnosed in stage I, 189 patients in stage II, and 81 patients in stage III (Figure...
2). As seen in figure 2, the risk of developing BC increases with age. BC is rare in females younger than 25 years. Approximately 6 to 7% of all breast cancer in the United are diagnosed in women under 40 years of age. As shown in figure 1, its prevalence gradually increases with age and reaches a plateau in women aged 46–65 years. Most BC is diagnosed in stage II (Figure 2).

Surgery type in relation to the stage of breast cancer

Surgery has been a mainstay of BC treatment for many years. In the case of the present study, the physicians performed different types of surgery (localized therapy), including lumpectomy, simple mastectomy, modified radical mastectomy, and other kinds of surgery to remove the breast tumor. Among the 334 BC patients who had surgery, sixty-six (66) patients underwent a lumpectomy, with a total of 22 patients in stage I, 36 in stage II, and 8 in stage III. Ninety-six (96) patients underwent a modified radical mastectomy, with a total of 8 patients in stage I, 49 in stage II, and 39 in stage III. Sixty-seven (67) patients underwent a simple mastectomy, with a total of 14 patients in stage I, 43 in stage II, and 10 in stage III. One hundred five (105) patients underwent other types of surgery with a total of 20 patients in stage I, 61 in stage II, and 24 in stage III (Figure 3). As seen in figure 3, most patients in stages II and III undergone modified radical mastectomy (a surgery technique that removes the entire breast including the breast tissue, skin, areola, and nipple) compared to simple mastectomy (a surgery technique that removes the breast including the breast tissue, nipple, areola, skin, and skin but not all the lymph nodes).

Survival rate in relation to the age

Age is a key risk factor for BC, and several scientific studies have suggested that patient age at diagnosis is associated with BC survival [33-36]. As seen in figure 4, young patients have a higher survival rate compared to older patients.

Survival rate in relation to breast cancer stages

Figure 5 presents the relative survival rates after surgery among 334 BC patients in relation to the stage. The total number of survivors detected among 334 BC patients diagnosed with invasive BC was 255 or 76.35%, including stage I, II, and III breast cancer. Sixty-six (19.76%) patients died after surgery and 13 (3.89%) of patients were not reported (there is no information available on whether those patients remained alive or dead).

Discussion

Breast Cancer (BC) is the most common cancer in women worldwide, with age considered to be an important risk factor [33]. Its incidence and prevalence rates continue to increase, especially in wealthy nations. Although it has been reported that 1 in 8 women will develop metastatic BC at some time in their life, a recent study has further stratified the risk of developing BC according to the age group, showing that the risk is 1 in 15,000 women up to age 25, 1 in 1900 up to age 30, and 1 in 200 women up to age 40 [37]. Similarly, our data demonstrated that the incidence and mortality rates of BC increase with age (Figure 1). The treatment of BC has traditionally employed a tri-modality approach: surgery followed by adjuvant chemotherapy and radiation therapy. Current BC treatment may include localized therapies such as surgery, cryotherapy, radiation therapy, chemical ablation, and/or systemic therapies such as chemotherapy, hormonal therapy, immune therapy, and targeted therapy used alone or in combination. In the case of the present study, the physicians performed different types of surgery (localized therapy), including lumpectomy, simple mastectomy, modified radical mastectomy, and others of surgery to remove the breast tumor. After the surgery, 255 out of 334 patients were alive, including patients with BC stages I, II, and III. Sixty-six (66) died, and the status of the 13 remaining patients was not reported (Figure 5). According to SEER Cancer Statistics Review, the overall 5-year relative survival for BC in females is 90 percent. For females diagnosed with stage I BC, the five-year relative survival is approximately 100%. However, it declines to 26% for females diagnosed with stage IV BC [38]. In addition to age and stage, other risk factors that influence BC survival include hormone receptor status, HER2 status, tumor grade, treatment, health insurance, and financial resources [39,40]. Furthermore, patient and physician factors such as attitudes, beliefs, and preferences influence treatment recommendations and delivery which are more likely to contribute to the survival differences [41].

Conclusion

The application of machine learning (ML) algorithms to the clinical dataset analyzed in this study revealed that among the 334 BC patients, 330 patients were women aged from 29 to 90 years old, and the remaining 4 patients were men aged from 44 to 84 years. Seventy percent (70%) of these patients were diagnosed with infiltrating ductal carcinoma, 27% with infiltrating lobular carcinoma, and 3% with mucinous carcinoma. These patients underwent surgery to remove the tumors. Among these individuals, 66 patients (22 in stage I, 36 in stage II, and 8 in stage III) underwent a lumpectomy; 96 patients (8 in stage I, 49 in stage II, and 39 in stage III) underwent a modified radical mastectomy; 67 patients (14 in stage I, 43 in stage II, and 10 in stage III) underwent a simple mastectomy; and 105 patients (20 in stage I, 61 in stage II, and 24 in stage III) underwent other types of surgery. After the surgery, 255 patients (252 women and 3 men) who underwent treatment for BC stages I, II, and III, remained alive, and sixty-six (65 women and 1 man) patients died. The status of the remaining 13 women patients was not reported. It is evident from our study that ML can help physicians to make quicker health decisions regarding the diagnosis, detection, prediction, and better treatment of BC patients.
Acknowledgment

This work was funded by the National Institutes of Health (NIH), Grant # 1U54MD015929-01 at Jackson State University, Jackson, MS, United States, and Faculty Research Award Program (FRAP) at Florida Agricultural and Mechanical University, Tallahassee, FL, United States.

Author contributions

Conceptualization, C.G.Y., S.S.T., and R.A.A.; Methodology, C.G.Y., J.G., and S.S.T; formal analysis, S.S.T., B.M., C.G.Y and R.E.; investigation, C.G.Y, S.S.T., K.J., and J.G.; supervision, K.S., and P.B.T.; 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.

Data availability statement

The breast cancer dataset that support the findings in this paper was made publicly available in Kaggle.com (https://www.kaggle.com/amandam1/breastcancerdataset).

References


