
THE ANALYSIS OF FACTORS AFFECTING STAGE LEVEL OF BREAST CANCER PATIENTS AT RSUD DR. SOETOMO SURABAYA

Destri Susilaningrum¹, Nur Azizah², Mutiah S³ and Mukti Ratna D⁴

Sepuluh Nopember Institute of Technology, Indonesia^{1,2,3 and 4}

Email: destr_its@yahoo.com¹, nur.azizah.1907@gmail.com²,
mutiahsalamah@yahoo.com³ and Mukti_Ratna@its.ac.id⁴

Keywords

Binary Logistic Regression, Breast Cancer, Stage

Article Info

Accepted:

December, 23th 2021

Revised:

January, 4th 2022

Approved:

January, 14th 2022

Abstract

Breast cancer is the cancer with the largest number in the world, and ranks as the fifth largest cause of death, which is 6.9% of the total 9,958,133 million cases of cancer deaths. Breast cancer is one of the main killers of women in the world and in Indonesia. Breast and cervical cancer dominate cancer cases in East Java. Throughout 2019, RSUD Dr. Soetomo Surabaya received 167,000 cancer patients, and the highest case was breast cancer. Binary logistic regression is a method of analysis that used to find out the relationship between the response variables (Binary or dichotomous) with the predictor variables that are polychotomous. The high number of breast cancer sufferers makes this disease requires special attention. The cancer staging or staging system is based on whether the cancer has spread from the breast to other parts of the body. Cancer stage is divided into two categories, namely early stage (stages 0 to III A) and late stage (stages III B to IV). This study will examine the factors that affect the level of cancer stage in breast cancer patients at RSUD Dr. Soetomo Surabaya in 2019. Using the binary logistic regression analysis method, it was found that the factors that significantly influence the stage of breast cancer are Grade and Obesity with a classification correctness level of 82.5%.

Introduction

Cancer is one of the main causes of morbidity and mortality in the world. According to the World Health Organization (WHO) in 2020 (Bray, Laversanne, Weiderpass, & Soerjomataram, 2021), breast cancer is the first cancer with the highest number in the world (Cao, Chen, Yu, Li, & Chen, 2021). Followed by lung cancer, colon cancer, prostate cancer, and stomach cancer (Cho, Park, & Kim, 2021). A survey conducted by WHO states that 45% of breast cancer incidences are in Asia (Sohn, Chang, & Miles, 2021).

Breast cancer also known as Carcinoma Mammae (Ca Mammae) is a malignant tumor that grows in breast tissue (Hwang, Peregrina, Maglalang, & Yoo, 2021). This cancer grows because of the abnormal growth of breast cells (Shaikh, Krishnan, & Thanki, 2021a). The cause of breast cancer is not known with certainty. So far it has been associated with female reproductive hormones (Bonfiglio & Di Pietro, 2021). This hormone is estrogen which plays a role in the growth and development of female reproductive organs, including breasts (Kumar et al., 2021). Breast cancer can spread or move to surrounding tissues, spread to lymph nodes (Shaikh, Krishnan, & Thanki, 2021b), enter blood vessels to other organs such as bones, lungs, liver, even the brain and cause malfunction of these organs which can lead to death (Mortezae, 2021). The stage

of breast cancer must be confirmed before the diagnosis is completed and treatment is selected (Metzger Filho et al., 2021). This process can determine whether the cancer has spread from the breast to other parts of the body. Cancer stages can be categorized into two, namely the Early stage which includes stages 0 to III A and the Last stage which consists of stages III B to IV.

RSUD Dr. Soetomo is a General hospital owned by the government of East Java Province, which is the largest hospital in East Java as well as a referral hospital for East Java and Eastern Indonesia. There is a poly that handles cancer, namely Poly Oncology where cancer treatment is carried out in an integrated manner and focused on organs, one of which is breast cancer. In 2019, Dr Soetomo Hospital was recorded 167,000 cancer patients, and the highest case was breast cancer. Cancer is also a high cause of death in Dr Soetomo Hospital Surabaya, from the 10 cancer cases hospitalized the highest are breast cancer patients, followed by cervical cancer and blood cancer.

This study aims to determine the factors that are thought to affect the level of breast cancer stage in breast cancer patients at Dr. Soetomo Hospital Surabaya in 2019. The analysis used was binary logistic regression. The benefit of this research is that knowing these causative factors, the patient can control these factors so that the development of staging can be controlled.

Research Method

Binary Logistic Regression

Binary logistic regression is a method of analysis that used to find out the relationship between the response variables (Binary or dichotomous) with the predictor variables that are polychotomous. The response variable (y) consists of two categories: "success" (y = 1) and "failure" (y = 0). The model logistic regression given as

$$\pi(\mathbf{x}_i) = \frac{\exp(\beta_0 + \beta_1 x_1 + \dots + \beta_j x_j + \dots + \beta_p x_p)}{1 + \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_j x_j + \dots + \beta_p x_p)} \quad (1)$$

The Equation 2.1 can be explained as a logit model: by logit transformation of $\pi(\mathbf{x}_i)$ as:

$$g(\mathbf{x}_i) = \ln\left(\frac{\pi(\mathbf{x}_i)}{1 - \pi(\mathbf{x}_i)}\right) \quad (2)$$

$$g(\mathbf{x}_i) = \beta_0 + \beta_1 x_{i1} + \dots + \beta_j x_{ij} + \dots + \beta_p x_{ip} = \sum_{j=0}^p \beta_j x_{ij}$$

The test is conducted to get the best model which was built by the significant parameters. Parameters were first tested simultaneously and then tested partially to get the significant parameters.

Simultaneously Test of Parameters

Hypothesis:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$$

$$H_1 : \text{at least one of } \beta_j \neq 0 ; j = 1, 2, \dots, p$$

Test Statistic:

$$G = -2 \ln \left[\frac{\left(\frac{n_1}{n}\right)^{n_1} \left(\frac{n_0}{n}\right)^{n_0}}{\sum_{i=1}^n \hat{\pi}_i^{y_i} (1 - \hat{\pi}_i)^{(1-y_i)}} \right] \quad (3)$$

With $n_1 = \sum_{i=1}^n y_i$; $n_0 = \sum_{i=1}^n (1 - y_i)$; $n = n_1 + n_0$

Reject H_0 if $G > \chi^2_{(\alpha; df)}$

Partial Test of Parameters

Hypothesis:

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0 ; j = 1, 2, \dots, p$$

Test Statistic:

$$W_j^2 = \frac{\hat{\beta}_j^2}{[SE(\hat{\beta}_j^2)]^2} \quad (4)$$

Reject H_0 if $W_j^2 > \chi^2_{(\alpha; df)}$

Goodness of Fit Model

The test for overall fit of a Binary logistic regression model with the hypothesis:

H_0 : Model Fit

H_1 : Model Not Fit

Test Statistic:

$$\hat{C} = \sum \frac{\left(\sum_{j=1}^{n_i} y_{ij} - \sum_{j=1}^{n_i} \hat{\pi}_{ij} \right)}{\left(\sum_{j=1}^{n_i} \hat{\pi}_{ij} \right) \left(1 - \frac{\sum_{j=1}^{n_i} \hat{\pi}_{ij}}{n_i} \right)} \quad (5)$$

y_{ij} : the Binary outcome for observation j in group i of the partition; $i = 1, \dots, g$; $j = 1, \dots, n_i$

$\hat{\pi}_{ij}$: the corresponding fitted probability for the model fitted to the ungrouped data.

Reject H_0 if $\hat{C} > \chi^2_{\alpha; (g-2)}$ or P-value $< \alpha$ or P-value $< \alpha$

Classification Procedure

This procedure was used to evaluate the result of prediction value given by the best model to compare with the observation value, give an evaluation on classification procedure to see the probability of miss classification. It is measured by *apparent error rate* (APER). APER value stated the proportion value of miss classification sample by the function of classification. If the subject only classified as two groups, y_1 and y_2 , then determination of classification errors can be known through the classification table described in Table 1.

Table 1. Classification Table.

Observation	Prediction	
	y_1	y_2
y_1	n_{11}	n_{12}
y_2	n_{21}	n_{22}

Description :

n_{11} : The number of the subject of y_1 on the correct classified as y_1

n_{12} : The number of the subject of y_1 on the miss classified as y_2

n_{21} : The number of the subject of y_2 on the miss classified as y_1

n_{22} : The number of the subject of y_2 on the correct classified as y_2

$$APER(\%) = \frac{n_{12} + n_{21}}{n_{11} + n_{12} + n_{21} + n_{22}} \times 100\% \quad (6)$$

And the proportion of the correct classification = $1 - APER(\%) = \left(\frac{n_{11} + n_{22}}{n_{11} + n_{12} + n_{21} + n_{22}} \right) 100\%$

Source of Data and Research Variables

This research used secondary data from the medical records of patients suffering breast cancer at RSUD Dr. Soetomo Surabaya in 2019 as many as 217 patients. The variables used consist of two kind variables: Respond variable, and Predictor variable.

Respond Variable

Respond Variable (Y) in this research is the level stage as given in Table 2.

Table 2. Respond Variable.

Variable	Level stage statue
Y	$Y = 1 =$ Early stage or $Y = 0 =$ Last stage

Predictors Variable (X)

Predictor variable (X) gives in Table 3.

Table 3. Research Variable.

Indicator	Variable	Categorical	Data Scale
X1	Grade	1= Grade 3 2= Grade 2 3= Grade 1	Ordinal
X2	Age	-	Ratio
X3	Obesity History	1= Obesity 2= No obesity	Nominal
X4	Comfort Level	1= Pain 2= No Pain	Nominal
X5	Patient's Psyche	1= Anxiety / Depression 2= Surrender	Nominal
X6	History of Anemia	1= Yes 2= No	Nominal
X7	Treatment	1= Operation (MRM, BCS) 2= Chemotherapy	Nominal
X8	Family History of Breast Cancer	1= Yes 2= No	Nominal
X9	Metastasis	1= Yes 2= No	Nominal

The operational definition of each variable is as follows.

A. Grade Breast cancer

The WHO using criteria, namely the Nottingham Grading system (also called Elston-Ellis modification of the Scarff-Bloom Richardson grading system). This rating scale looks at 3 different cell images and is classified each score from 1-3. The classifications are:

Table 4. Breast Cancer Grade

Grade	Score	Information
Grade I	Score 3-5	Low grade with well differentiated cancer (well differentiated) where the cancer cells do not grow quickly and do not appear to spread.
Grade II	Score 6-7	Cancer with moderate differentiation (moderately/intermediate differentiated) which has a picture between grades 1 and 3.
Grade III	Score 8-9	Cancer that is poorly differentiated (poorly differentiated or undifferentiated) in which cancer cells grow rapidly and are more likely to spread.

Another benefit of determining the grading is to determine the response to therapy that will be given (Crotty et al., 2021). At a poor degree of differentiation, where cell growth and spread is considered to be more rapid or aggressive (Redmond, McCarthy, Buchanan, Levingstone, & Dunne, 2021), additional therapy is needed besides definitive, namely by administering chemoradiation. The relationship between stage and grade of breast cancer is that the stage in breast cancer serves to describe the condition of the cancer, namely the location and spread of cancer cells and how far it affects other body organs.

B. Comfort Level

Pain is a common complaint after treatment of breast cancer patients, even years after treatment (De Baets et al., 2021). Cancer pain is often found in patients who first come for treatment (Drewes, Kuhlman, Andresen, & Olesen, 2021).

C. Patient’s Psyche

Reactions in some people with cancer vary widely, such as shock, fear, anxiety, feelings of grief, anger, sadness, and even withdrawal. Anxiety is a mental condition filled with worry and fear of what might happen, both related to limited problems and strange things.

D. Anemia

Anemia in cancer patients can cause fatigue and decreased quality of life and increased mortality. In general, the presence of anemia in cancer patients can increase mortality by up to 65%. Patients are said to be anemic if they have hemoglobin < 12 g/dL.

E. Family History of Breast Cancer

Family history of inherited breast cancer is a risk factor for breast cancer. Family history of breast cancer will increase the development of breast cancer at a young age.

F. Treatment

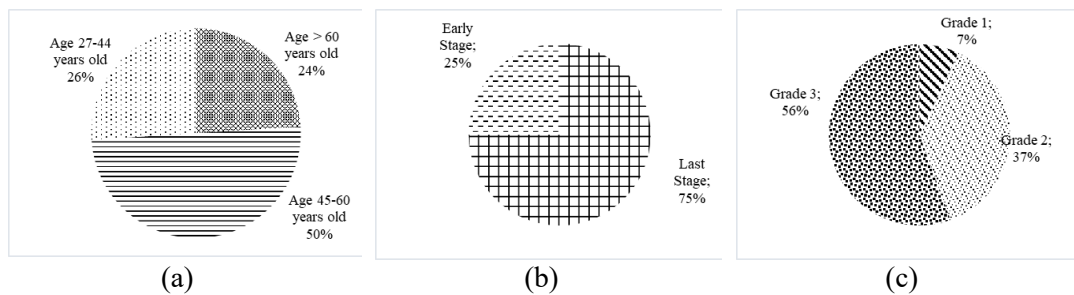
Breast cancer treatment consists of surgery, radiation therapy, chemotherapy and hormone blocking drugs.

G. Metastases

Cancer cells can infiltrate the surrounding tissues and spread (metastasize) through blood vessels and lymph vessels. Metastases are cancer cells that have spread beyond the organs or tissues from where the cancer first appeared. Spread of cancer cells to other parts of the body is affected by many things, such as the type of cancer, the severity of cancer stage, and the location of the cancer originating [15].

Result and Discussion

The Characteristics of Breast Cancer Patient at RSUD Dr Soetomo



Picture 1. the characteristic of (a) The Age, (b) The Stage Level, and (c) The Grade Level.

Picture 1. a) show that the Age with high percentage of sufferer breast cancer at RSUD Dr. Soetomo Surabaya are the ages between 45 to 60 years old about 50% of the breast cancer patients, while between the ages of 27 - 44 and more than 60 years each with almost the same percentage, namely 26% and 25%.

Picture 1. b) show that majority the breast cancer patients at RSUD Dr. Soetomo Surabaya in Last stage condition about 75%. And figure 1 c) show that the greater Grade is Grade 3 (Cancer that is poorly differentiated) about 56%, than Grade 2 with 37% and the rest are Grade 1 (7%).

The Factors That Affect the Stage Level of Breast Cancer

The result of the binary logistic regression analysis, with the full logit model which include all variables gives as below.

$$g(x_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_9 x_{i9} = \sum_{j=1}^9 \beta_j x_{ij}$$

Simultaneously Test of Parameter

To find out which parameters are significant, the simultaneously test with the hypothesis as in subsection 2.1.1 had done and give the result on Table 5.

Table 5. Simultaneous Test Results.

χ^2	Df	P-value
85.615	9	0.000

For $\alpha = 0.10$, it showed that P-value = 0.000 less than α indicated that reject H_0 , the partial test was then conducted to find out the variables that significant in the model.

Partially Test of Parameter

The partial test with the hypothesis as in subsection 2.1.2, with the significant level 10%, give the result of the significant variables as shown in Table 6.

Table 6. Partial Test Results.

Variables	B	Wald	Df	P-value
X ₁ : Grade	1.303	14.877	1	0.000
X ₃ : Obesity History	-0.760	3.244	1	0.072
Constant	-1.428	5.285	1	.022

Table 6 shows that the P-value less than $\alpha = 0.10$, this indicate the variables are significant to the model, and the variables are: Grade and Obesity History. The logit Binary logistics Regression model gives as below.

$$g(x) = -1.428 + 1.303x_1 - 0.76x_3$$

3.2.3 Goodness of fit model

The hypothesis of goodness of fit model as in subsections 2.1.3, shown in Table 7.

Table 7. Goodness Fit of Model Results.

χ^2	Df	P-value
5.935	8	0.654

Table 7. showed that P-value is more than $\alpha = 0.10$, indicated the test cannot reject H₀, and so the model fit.

Classification Procedure

The subject only classified as two groups, with Y₁: Secure and Y₂: Insecure. The determination of classification errors can be known through the classification table described in Table 8.

Table 8. Classification Table.

Observed	Prediction		Percentage Corrected
	Early Stage	Last Stage	
Early Stage	25	30	45.5
Last Stage	8	154	95.1
	Overall Percentage		82.5

Based on the results above, the factors affecting stage level of breast cancer patients at RSUD Dr. Soetomo Surabaya are Grade and Obesity history with the correct classification between predicted and observation 82.5%.

Conclusion

The factors that affect the level stage of breast cancer patients at RSUD Dr. Soetomo Surabaya are Grade and Obesity History.

References

Bonfiglio, R., & Di Pietro, M. L. (2021). The impact of oral contraceptive use on breast cancer risk: State of the art and future perspectives in the era of 4P medicine. *Seminars in Cancer Biology*. Elsevier.

- Bray, Freddie, Laversanne, Mathieu, Weiderpass, Elisabete, & Soerjomataram, Isabelle. (2021). The ever-increasing importance of cancer as a leading cause of premature death worldwide. *Cancer*.
- Cao, Wei, Chen, Hong Da, Yu, Yi Wen, Li, Ni, & Chen, Wan Qing. (2021). Changing profiles of cancer burden worldwide and in China: a secondary analysis of the global cancer statistics 2020. *Chinese Medical Journal*, 134(7), 783.
- Cho, Do-Yeon, Park, Juhee, & Kim, Dong-Sook. (2021). The impact of expanding health insurance coverage for anti-cancer drugs on cancer survival in Korea. *Cancer Medicine*.
- Crotty, Erin, Downey, Kira, Ferrerosa, Lauren, Flores, Catherine, Hegde, Bindu, Raskin, Scott, Hwang, Eugene, Vitanza, Nicholas, & Okada, Hideho. (2021). Considerations when treating high-grade pediatric glioma patients with immunotherapy. *Expert Review of Neurotherapeutics*, 21(2), 205–219.
- De Baets, Liesbet, Devoogdt, Nele, Haenen, Vincent, Evenepoel, Margaux, Dams, Lore, Smeets, Ann, Neven, Patrick, Geraerts, Inge, De Vrieze, Tessa, & De Groef, An. (2021). Cognitions and physical impairments in relation to upper limb function in women with pain and myofascial dysfunctions in the late stage after breast cancer surgery: an exploratory cross-sectional study. *Disability and Rehabilitation*, 1–8.
- Drewes, Asbjørn Mohr, Kuhlman, Louise, Andresen, Trine, & Olesen, Søren Schou. (2021). Medical Treatment of Pain in Chronic Pancreatitis: Guidelines for Clinical Practice. *Clinical Pancreatology for Practising Gastroenterologists and Surgeons*, 273–282.
- Hwang, Jane, Peregrina, Hillary Nicole, Maglalang, Dale Dagar, & Yoo, Grace J. (2021). *A qualitative exploration of the trajectory of help seeking among Asian American breast cancer survivors*.
- Kumar, Nitish, Gulati, Harmandeep Kaur, Sharma, Aakriti, Heer, Shilpa, Jassal, Anupmjit Kaur, Arora, Lovenish, Kaur, Simranpreet, Singh, Atamjit, Bhagat, Kavita, & Kaur, Arshmeet. (2021). Most recent strategies targeting estrogen receptor alpha for the treatment of breast cancer. *Molecular Diversity*, 25(1), 603–624.
- Metzger Filho, Otto, Viale, Giuseppe, Stein, Shayna, Trippa, Lorenzo, Yardley, Denise A., Mayer, Ingrid A., Abramson, Vandana G., Arteaga, Carlos L., Spring, Laura M., & Waks, Adrienne G. (2021). Impact of HER2 heterogeneity on treatment response of early-stage HER2-positive breast cancer: phase II neoadjuvant clinical trial of T-DM1 combined with pertuzumab. *Cancer Discovery*.
- Mortezaee, Keywan. (2021). Organ tropism in solid tumor metastasis: an updated review. *Future Oncology*, 17(15), 1943–1961.
- Redmond, John, McCarthy, Helen, Buchanan, Paul, Levingstone, Tanya J., & Dunne, Nicholas J. (2021). Advances in biofabrication techniques for collagen-based 3D in vitro culture models for breast cancer research. *Materials Science and Engineering: C*, 122, 111944.
- Shaikh, Khalid, Krishnan, Sabitha, & Thanki, Rohit. (2021a). *Artificial Intelligence in Breast Cancer Early Detection and Diagnosis*. Springer.
- Shaikh, Khalid, Krishnan, Sabitha, & Thanki, Rohit. (2021b). Types, Diagnosis, and Treatment of Breast Cancer. In *Artificial Intelligence in Breast Cancer Early Detection and Diagnosis* (pp. 21–35). Springer.
- Sohn, Young Jin, Chang, Connie Y., & Miles, Randy C. (2021). Current Gaps in Breast Cancer Screening Among Asian and Asian American Women in the United States. *Journal of the American College of Radiology*.

Copyright holder :

**Destri Susilaningrum¹, Nur Azizah², Mutiah S³, Mukti Ratna
D⁴ (2021)**

First publication right :

Devotion : Journal of Research and Community Service

This article is licensed under:

