Logistic Regression Model of Relationship between Breast Cancer Pathology Diagnosis with Metastasis

M N Bustan\(^1\)\(^*\) and B Poerwanto\(^2\)

\(^1\)Statistics Study Program, Universitas Negeri Makassar, Makassar
\(^2\)Department of Informatics Engineering, Universitas Cokroaminoto, Palopo

\(^*\)e-mail: mnbustan@unm.ac.id

Abstract. Management of breast cancer requires a diagnosis of breast cancer, including an accurate pathology diagnosis to direct the provision of appropriate treatment. Part of inseparable diagnosis of breast cancer itself is determining the presence of metastasis. This study aims to determine the relationship model of pathological diagnosis with the possibility of metastasis to other organs of the breast. Pathology diagnosis results include tumor position, topography, behavior, grade, and tumor status. This study used secondary data from the Hospital Information System (SIRS) of the hospital in Makassar, Indonesia. Data obtained from the results of pathology and clinical examinations of breast tumor as many as 204 patients were hospitalized during the year of 2018. Binary regression analysis modeling was performed to see the relationship between types of pathological diagnosis (laterality, topography, behavior, grade, tumor status) of breast cancer with the possibility of metastases to other organs. The results of the analysis found a significant relationship between the position and status of the tumor with the presence of metastasis. The results of the metastatic binary logistic regression model was

\[
\log \left( \frac{p}{1-p} \right) = \beta_0 + \beta_1 \text{laterality} + \beta_2 \text{topography} + \beta_3 \text{behavior} + \beta_4 \text{grade} + \beta_5 \text{tumor status}
\]

Chi-square analysis between metastasis with surgery and chemotherapy found the significant relationships between metastasis with surgery \((X^2 = 15.951; \ p \text{ value} = 0.001)\) and metastasis with chemotherapy \((X^2 = 20.796; \ p \text{ value} = 0.001)\). The results of this analysis indicate that the position and status of the tumor have a relationship with the likelihood of metastasis. In addition, this metastasis further determines the management of treatment whether surgery or chemotherapy or both.

Keywords: Breast cancer, pathological diagnosis, metastasis, logistic regression

1. Introduction
Breast cancer is a systemic disease or local disease in the breast that has the potential to spread or metastasize to other organs.[1], [2] Breast cancer is the most common cancer among women, and is one of the 10 main causes of death.[3] Therefore, prevention and control of breast cancer is very important to reduce the incidence and increase survival rates for breast cancer patients.[4] To prevent and control breast cancer, early detection and accurate diagnosis are needed.[5] The diagnosis of the presence of breast cancer is very complex, at least including clinical diagnosis, pathological diagnosis, and radiological diagnosis. [6],[7]

Breast cancer management requires an accurate diagnosis in order to get the right treatment options.[8] To carry out a prompt treatment required a complete diagnosis that comes from various types of examinations. Type of breast cancer diagnostic examination consists of clinical
examination (location and status of tumor), anatomical or cytology examination (topography, behavior, morphology, grade), histopathology, molecular (immunohistology, genetic), and radiological examination (mammography, ultrasound, MRI, and tomography). [9] [10] [11] [12] Information from the results of the diagnostic examination is used to direct the type of surgery to be performed or what type of drug should be given. [9], [13]

In addition, determining the presence of metastasis is an inseparable part of the diagnosis. [10] This metastasis can spread to the lymph glands and other distant organs such as the liver, brain, and bones. [1], [13]. Breast cancer is a type of cancer that is more potentially metastasis.[14] The presence of metastasis is the important diagnosis in determining the direction of surgery and /or chemotherapy to be given. [15]

Metastasis is an important factor in cancer development because metastasis is responsible for most cancer deaths.[16] In the most cases, cancer patients with local tumors have a better chance of survival compared to metastatic tumors.[17]. The presence of breast cancer metastases is related to cancer status (benign or malignant), cell morphologic abnormalities, cancer cell behavior, and the degree of malignancy.[18], [19] This condition can be found based on the results of histopathological examination. [20] [21]

Based on the importance of determining the diagnosis of breast cancer and the presence of metastases, this study was conducted to find a model of the relationship between anatomic pathological diagnosis and the possibility of metastasis. The statistical analysis model used was binary logistic regression to determine the type of pathological diagnosis associated with the occurrence of metastasis. [22] In this study binary logistic regression was used as a statistical analysis model that explains the relationship between the binary response variable Y (metastasis) with several independent variables Xn (topography, behavior, grade, tumor type). [23] Then the relationship analysis was done by using the Chi-square test to see the relationship of this metastasis with surgery and / or chemotherapy given by the doctor.

2. Material and Methods
A total of 204 breast cancer patient data had been collected from the results of anatomic pathology examination and hospital information system (SIRS) during the year of 2018. These data consisted of many variables and the one that were used directly in this study were: tumor position (laterality), topography, behavior, grade, and tumor status. This variable grouping was adjusted to the classification and code of ICD X, ICD Oncology, and Canreg5. Laterality variable was grouped into right, left, and bilateral; topography consists of codes C50.0 through C50.9; behavior includes benign, uncertain, carcinoma in-situ, malignant; grade included well, moderate, and poor differentiated; type of tumor status divided into benign and malignant.

To the complete data that has been available with the metastasis response variable along with the predictor variables, a binary logistic regression model analysis was performed.

The binary logistic regression model that is distributed by Bernoulli is

$$\pi(x) = \frac{e^{(\beta_0 + \beta_1x_1 + ... + \beta_px_p)}}{1 + e^{(\beta_0 + \beta_1x_1 + ... + \beta_px_p)}}$$

where the value \( \pi(x) \) is the chance of a successful occurrence, namely \( y = 1 \), while \( \beta_p \) is the parameter value for the value \( j = 0,1,2, ..., p \). For parameter estimation, this logistic regression model is changed through logit transformation to

$$g(x) = \ln \left( \frac{\pi(x)}{1 - \pi(x)} \right) = \beta_0 + \beta_1x_1 + ... + \beta_px_p.$$  

Then the parameters were assessed using the Maximum Likelihood Estimation (MLE) method. To obtain the estimated parameters the iteration process was carried out by the Newton-Raphson method, by determining the initial value of \( \beta \), namely \( \beta^0 \).

The parameters of the model are estimated from vector \( \beta^T = (\beta_0, \beta_1, \beta_2, ..., \beta_p) \)

After estimating the parameters obtained, then testing the significance of the existing parameters in the model. Testing the significance of the parameters was done by doing a test simultaneously using the test statistic used is the Likelihood Ratio Test (\( G^2 \)).
$$G = 2 \left\{ \sum_{j=1}^{p} \left[ y_j \ln(\hat{\pi}_j) + (1 - y_j) \ln(1 - \hat{\pi}_j) \right] - [n_1 \ln(n_1) + n_0 \ln(n_0) - n \ln(n)] \right\}$$

Partial test to determine the significance of each parameter in the model was using the Wald test. This test was conducted to determine whether each variable can be relied upon to build a model or not in the classification process. The Wald test formula was

$$W = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)}$$

To find out whether the model with the predictor variable was an appropriate model, it was necessary to do a model suitability test using the Chi-Square test statistics, where

$$\chi^2 = \sum_{k=1}^{g} \left( \frac{O_k - n_k\bar{\pi}_k}{n_k\bar{\pi}_k(1 - \bar{\pi}_k)} \right)^2$$

Interpretation of the coefficient of this parameter was done to determine the tendency or functional relationship between the predictor variable and the response variable, and showed the effect of changes in value on the variable concerned. Odds ratio, denoted $\varphi$, was defined as the ratio between the two logit values at $x = 1$ and $x = 0$, then:

$$\varphi = \frac{\pi(1)/1 - \pi(1)}{\pi(0)/1 - \pi(0)}$$

3. Results

Binary logistic regression modeling had been carried out between factors of tumor position, tumor status, behavior, and grade to the presence of metastatic breast cancer. The results of this study indicate that the position of the cancer (left, right, bilateral) and tumor status (benign, malignant) is related to the presence of metastasis. As the result of the binary logistic regression model of metastatic breast cancer, it is obtained logistic regression parameter values from the position and status of the tumor in the equation of the chance of a metastasis as follows:

$$\pi(x) = \frac{\exp(1.330 - 0.670lat1 - 0.779lat2 - 0.724status)}{1 + \exp(1.330 - 0.670lat1 - 0.779lat2 - 0.724status)}$$

For parameter estimation, a logistic regression logit model is written in the equation:

$$g(x) = 1.330 - 0.670lat1 - 0.779lat2 - 0.724status$$

The results of this model show that tumor position factors are associated with the presence of metastases where bilateral tumor positions have a greater likelihood of metastasis compared to those left or right.

The odds ratio value for each left and right position is Odds Ratio left = 0.51 (95% CI 0.23-1.06); Odds Ratio right = 0.46 (0.24-0.90). Concerning the status of the tumor with metastasis, the analysis found that malignant tumors is more likely to be metastatic compared with benign tumors with benign Odds Ratio benign = 0.49 (95% CI 0.23 - 1.02). The magnitude of the influence of these two predictors on the incidence of metastasis is 7.1% based on the 0.071 Nagelkerke Pseudo R2 value.

Metastasis is an inseparable part of the diagnosis and has an interest in making the right decision or treatment. Chi-square analysis between metastases with surgery and chemotherapy found a significant relationship between metastasis with surgery ($\chi^2 = 15.951$; $p$ value = 0.001) and metastasis with chemotherapy ($\chi^2 = 20,796; p$ value = 0.001).

4. Discussion

Laterality factors, tumor status, grade, behavior are indeed factors that contribute to the likelihood of tumor spread to external organs of the breast. [10], [19], [24] Breast cancer is more likely to attack the left breast than the right. [17] However, it is unclear to what position the possibility of metastasis can occur. [25]
This study also does not indicate whether the metastasis is more often derived from left or right breast cancer, but states that if breast cancer is associated with bilateral breast cancer that is more likely to have metastasis.

In terms of tumor status, malignant tumors are more likely to spread compared to benign. [21] This study find the same thing where malignant tumors are more metastatic compared to benign tumors.

The degree of malignancy or grade is related to the possibility of metastasis. The higher the grade the more likely it is to have metastasis.[26]

This research both in partial test and overall test did not find a relationship between grade and metastasis. Tumor position on the breast or topography may also be associated with metastasis. [27] This study find no difference in the tumor site, whether in certain parts of the breast, or in some parts of the breast (multiple) or unclear position with the spreading of breast cancer.

Behavior or how the histopathological picture of the cell enlargement is expected to be possibly related to metastasis. [28] This study find no difference in behavior, whether benign, carcinoma insitu, or malignant with the spread of breast cancer to other organs. Therefore, this binary logistic regression model is unable to find the relationship between pathological diagnosis factors (topography, behavior, and grade) with the spread of breast cancer to other organ distant, except in relation to laterality and tumor status.

5. Conclusions
The conclusion of this study shows that the position and status of the tumor have a relationship with the likelihood of metastasis. In addition, this metastasis further determines the direction of treatment whether surgery or chemotherapy or both.

Acknowledgments
Thank you for the financial support from the Directorate General of Research and Development Strengthening (DRPM) of the Ministry of Research, Technology and Higher Education (Kemenristekdikti) and the Institute of Research and Community Service (LPPM) Universitas Negeri Makassar for all administrative and technical assistance in carrying out this research.

References
[1] Redig AJ, McAllister SS. Breast cancer as a systemic disease: a view of metastasis. J Intern Med 2013; 274: 113–26.
[2] Weigelt B, Peterse JL, van’t Veer LJ. Breast cancer metastasis: markers and models. Nat Rev Cancer 2005; 5: 591–602.
[3] Ma L. Determinants of breast cancer progression. Sci Transl Med 2014; 6: 243fs25.
[4] Giordano SH, Buzdar AU, Smith TL, et al. Is breast cancer survival improving? Cancer 2004; 100: 44–52.
[5] Montella M, Crispo A, D’Aiuto G, et al. Determinant factors for diagnostic delay in operable breast cancer patients. Eur J Cancer Prev 2001; 10: 53–9.
[6] McDonald ES, Clark AS, Tchou J, et al. Clinical Diagnosis and Management of Breast Cancer. J Nucl Med 2016; 57 Suppl 1: 9S-16S.
[7] Pathology JHU. Types of Breast Cancer - Breast Cancer. Johns Hopkins Medicinehttps://pathology.jhu.edu/breast/my-results/types-of-breast-cancer (2018, accessed September 19, 2018).
[8] Sharma GN, Dave R, Sanadya J, et al. Various types and management of breast cancer: an overview. J Adv Pharm Technol Res 2010; 1: 109–26.
[9] Nounou MI, ElAmrawy F, Ahmed N, et al. Breast Cancer: Conventional Diagnosis and Treatment Modalities and Recent Patents and Technologies. Breast Cancer (Auckl) 2015; 9: 17–34.
[10] Lee AHS. The histological diagnosis of metastases to the breast from extramammary malignancies. J Clin Pathol 2007; 60: 1333–41.
[11] Orel SG, Schnall MD. MR Imaging of the Breast for the Detection, Diagnosis, and Staging
of Breast Cancer. Radiology 2001; 220: 13–30.

[12] Singh AK, Gupta B. A Novel Approach for Breast Cancer Detection and Segmentation in a Mammogram. Procedia Comput Sci 2015; 54: 676–682.

[13] Abner AL, Collins L, Peiro G, et al. Correlation of tumor size and axillary lymph node involvement with prognosis in patients with T1 breast carcinoma. Cancer 1998; 83: 2502–2508.

[14] Kuru B, Camlibel M, Dinc S, et al. Prognostic factors for survival in breast cancer patients who developed distant metastasis subsequent to definitive surgery. Singapore Med J 2008; 49: 904–11.

[15] Kim KJ, Huh SJ, Yang J-H, et al. Treatment Results and Prognostic Factors of Early Breast Cancer Treated with a Breast Conserving Operation and Radiotherapy. DOI: 10.1093/jjco/hyi039.

[16] Afef Khanfira Faiez Lahiania RacemBouzguenda InesAyediaJamelDaoudb MounirFrikhaa. Prognostic factors and survival in metastatic breast cancer: A single institution experience. Reports Pract Oncol Radiother 2013; 18: 127–132.

[17] Bao J, Yu K-D, Jiang Y-Z, et al. The effect of laterality and primary tumor site on cancer-specific mortality in breast cancer: a SEER population-based study. PLoS One 2014; 9: e94815.

[18] Johns Hopkins Medicine. Malignant Tumors - An Atlas of Breast Imageshttps://www.hopkinsmedicine.org/breast_center/treatments_services/breast_cancer_diagnosis/breast_pathology/malignant_tumors.html (2018, accessed October 18, 2018).

[19] Leong AS-Y, Zhuang Z. The changing role of pathology in breast cancer diagnosis and treatment. Pathobiology 2011; 78: 99–114.

[20] Lundin M, Lundin J, Helin H, et al. A digital atlas of breast histopathology: an application of web based virtual microscopy. J Clin Pathol 2004; 57: 1288–91.

[21] Sharma K, Vyas SP, Dhayal S. Clinical and histopathological correlation of breast lesions. Int J Res Med Sci 2018; 6: 1348.

[22] David G. Kleinbaum MK. Logistic Regression: A Self-Learning Text - David G. Kleinbaum, Mitchel Klein - Google Books.

[23] Bustan MN, Tiro MA, Annas S, et al. Analysis of Ordinal Logistic Regression Model on Breast Cancer Diagnosis by Birads Mammography. Indian J Public Heal Res Dev 2019; 10: 1199–1203.

[24] Sinn H-P, Kreipe H. A Brief Overview of the WHO Classification of Breast Tumors, 4th Edition, Focusing on Issues and Updates from the 3rd Edition. Breast Care 2013; 8: 149–154.

[25] Amer MH. Genetic factors and breast cancer laterality. Cancer Manag Reg Sci 2014; 6: 191–203.

[26] Elston CW. Histological Grading of Breast Cancerhttp://tvmouse.ucdavis.edu/bcancercd/311/grading_diagram.html (accessed September 26, 2017).

[27] Gerard JP, Noël P, Mayer M, et al. A topographic approach to breast cancer: the relation of topographic and mammographic findings. Cancer 1977; 40: 928–30.

[28] Gotzsche PC, Jørgensen KJ. Screening for breast cancer with mammography. Cochrane Database Syst Rev. Epub ahead of print June 4, 2013. DOI: 10.1002/14651858.CD001877.pub5.