Correspondence Analysis of Breast Cancer Diagnosis Classification

by Suwardi Annas

Submission date: 23-Dec-2020 02:40AM (UTC-0500)
Submission ID: 1480800723
File name: 7._Nadjib_Bustan_2019_J._Phys.__Conf._Ser._1244_012030.docx (1.43M)
Word count: 2869
Character count: 16209

Correspondence Analysis of Breast Cancer Diagnosis Classification

M. Nadjib Bustan*, M. Arif Tiro, Suwardi Annas
Department of Statistics, Faculty of Mathematics and Natural Science, Universitas Negeri Makassar, South Sulawesi, Indonesia

*Corresponding author: mnbustan@unm.ac.id

Abstract. There are five indicators used in the diagnosis of breast cancer classification. The indicators are a type of malignancy, location, topography, morphology, behavior, and grade. This study aimed to assess how the relationship between types of diagnostic classification was given to breast cancer patients. The research is a quantitative method and used the hospital medical records were collected in the form of anatomical pathology examination results for hospital patients in the year 2017. Data were obtained from 317 pathology examinations which included 282 breast cancer patients. Each patient is given a diagnosis according to the pathologist's observations into various classifications according to location, topography, type, morphology, grade, and behavior. The result of the analysis showed a relationship between the location of cancer and the type of malignancy. Furthermore, there is a difference in the probability of a malignant or benign tumor-attacking the right or left breast. Correspondence analysis was carried out between the location of the tumor with topography, type of malignancy, morphology, grade, and behavior respectively. The results showed that there was a significant correspondence between topography with the type of malignancy, type of malignancy with morphology, morphology with grade, and grade with behavior. Each type of diagnosis of breast cancer diagnosis is entirely accurate and has a significant correspondence relationship with each other.

Keywords: breast cancer diagnosis classification, topography, morphology, grade, behavior

1. Introduction

Determining the right and precise diagnosis is critical to determine the direction of action or treatment of breast cancer. The right diagnosis of breast cancer is a gradual process, starting from the clinical diagnosis by the surgeon, then a radiological diagnosis, and followed by a pathology diagnosis [1]. The results of the pathology examination will be the basis for making appropriate therapeutic decisions. So that, the investigation of pathology examination should be completed and clearly. The diagnosis of this pathology includes four types, namely topography, morphology, behavior, and grade [2].

The type of pathological abnormalities that occur in breast cancer is related to the anatomical structure of the breast itself which consists of epithelial tissue and supporting tissue (connective tissue). Breast epithelial tissue is lobules and ducts which when attacked by a tumor tend to become malignant tumors. The fat tissue (*adipose tissue*) and fibrous tissue (*fibrous tissue*) can develop into benign tumors such as *fibroadenomas* and *sarcoma* [3].

The classification of each method of diagnosis is needed in determining the treatment choices made. Topographic diagnosis classification shows the anatomical position of the tumor. Classification or M code diagnosis of morphology is based on a description of abnormalities or damage to the shape of breast cells that are useful in directing treatment or chemotherapy that is suitable for administration. Behavioral diagnosis is also based on the description of abnormalities or damage to the shape of breast cells that classify benign upper breast cancer, in-situ carcinoma, malignancy, and metastases. This classification is used to help doctors determine the feasibility of surgery. Grade classification is useful for predicting breast cancer prognosis [4].

As measuring the accuracy of diagnosis, correspondence analysis can be done by analyzing the relationship between two types of diagnostic classification or between one type of diagnosis and another diagnosis [5]. Correspondence analysis is part of the multivariate analysis that studies the relationship between two or more variables by displaying rows and columns together from the contingency table [6]. Correspondence analysis is one of the descriptive statistical methods designed to analyze two-way or multi-direction contingency tables that contain relationships. Between row and column variables [7]. The results of the correspondence analysis show the best dimensions for presenting data in the form of perception maps. Contingency table preparation will be filled with two types of diagnosis of breast cancer diagnosis. The chi-square value explains the size of the proximity of each category variable. The accurate diagnostic model is needed that can be developed through the development and testing of statistical models, such as genetic algorithms, regression analysis, especially logistic regression (binary, ordinal, multinomial), artificial neural networks, and gene expression signatures [8]–[10]. One of them is the analysis of correspondence.

The development of this correspondence analysis model is based on input from various types of breast cancer diagnosis classifications [2]. This study assesses how the relationship among several types of diagnosis classification is given to breast cancer patients by using correspondence analysis.

2. Materials and Methods

Hospital medical records were collected in the form of anatomical pathology examination results for all hospital patients during the year 2017. Data were obtained from 317 pathology examinations which included 282 breast cancer patients. Each patient was given a diagnosis according to the pathologist's observations into various classifications according to location, topography, type, morphology, grade, and behavior. All of this diagnosis classification were considered as the categorical data. Type of location and topography were classified into nominal variables. Behavioral classification and grade were ordinal variables. Data analysis using the quantitative descriptive method to describe frequency and percentage. The correlational analysis is using the SPSS 12.0 software. Relationship analysis was carried out between locations with other diagnosis classifications. Proximity analysis was made with the framework of the series relationship between tumor location, topography, type of tumor, morphology, grade, and behavior (Figure 1).

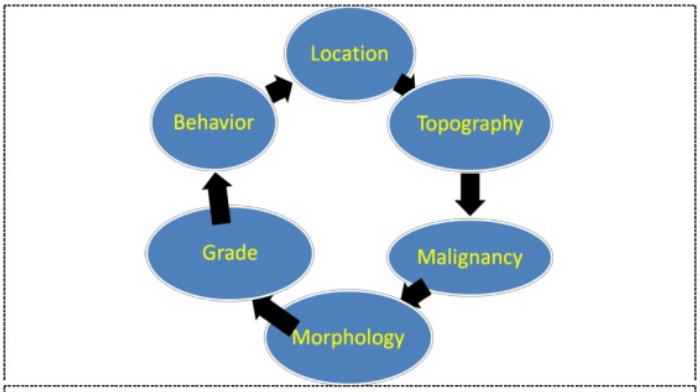

Figure 1. Correspondence cycle of breast cancer diagnosis

3. Results

One indication of cancer is the presence of lumps in the breast of the sufferer. The result of observation of lumps location is divided into three parts, namely left, bilateral and right. Results of distribution lumps location are presented in Table 1.

Table 1. Frequency distribution of breast cancer diagnosis by the location of lumps

Location of lumps	Frequency	Percent
Left	146	51.8
Right	108	38.3
Bilateral	28	9.9
Total	282	100.0

Table 1 shows that most lumps are located on the right side of the breast which is 51.8%. Furthermore, the results of data analysis showed that 38.3% of lumps were found on the right side of the breast and only a small proportion of tumors had bilateral positions. The type or degree of malignancy can be broadly divided into malignant tumors and benign tumors (Figure 2). Type of malignancy distribution is presented in Table 2. The results showed that most tumors included type malignancy (81.6%).

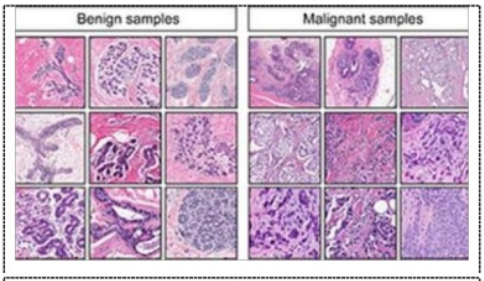

Figure 2. Pathological overview of breast cancer according to the type ofmalignancy [11]

Table 2. Frequency distribution of breast cancer diagnosis by type of malignancy

Type of Malignancy	Frequency	Percent
Malignant	230	81.6
Benign	52	18.4
Total	282	100.0

Topography classified according to the quadrant position. The quadrant started from the nipple, the quadrant over the left and right (upper inner-outer), and the quadrant below the left and rght (inferior inner-outer) (Figure 3). Table 3 represented that the distribution of topography is divided into seven group.

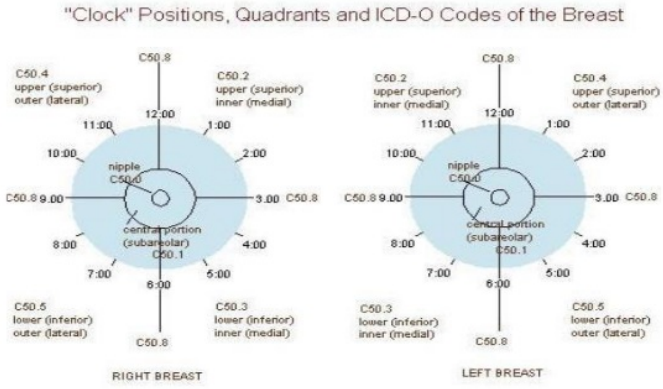

Figure 3. Types of breast cancer according to topography [12]

Table 3. Frequency distribution of breast cancer diagnosis by topography

Topography	Frequency	Percent
Nipple	1	0.4
Upper outer	1	0.4
Lower Outer	1	0.4
Axilla tail	1	0.4
Multiple	217	77.0
Benign	57	20.1
Others	4	1.3
Total	282	100.0

Table 3 describes that most breast cancers are multiple topography. The form of multiple cancer indicated that are more than one lumps and spread in the breast. The number of diagnoses of cancer located in the nipple, upper outer, lower outer, and axilla tail is indicated by a very small value. Classification of morphology divides into eight types of breast cancer abnormalities. This type is based on changes in cell image and tissue structure (Figure 4).

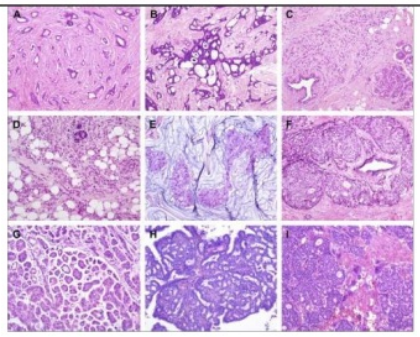

Figure 4. Morphology classification of breast cancer [13]

Table 4. Frequency distribution of breast cancer diagnosis by morphology

Morphology	Frequency	Percent
Neoplasm	24	8.5
Epithelial tumor	4	1.4
Intraductal carcinoma	167	59.2
Fibroadenoma	51	18.1
Adenofibroma	3	1.1
Phyllodes tumor	11	3.9
Medullary carcinoma	3	1.1
Others	19	6.7
Total	282	100.0

Table 4 describes that morphology of breast cancer as *intraductal carcinoma* (59.2%). Furthermore, there are 18.1% of cases of breast cancer in the form of *fibroadenoma* (18.1%). Amount of data for diagnosis of breast cancer with morphology neoplasm as much as 8.5%. The variable behavior of breast cancer covered four types that benign, *in situ carcinoma* and *malignant*. Figure 5 shows the form of breast cancer with *lobular carcinoma* in situ.

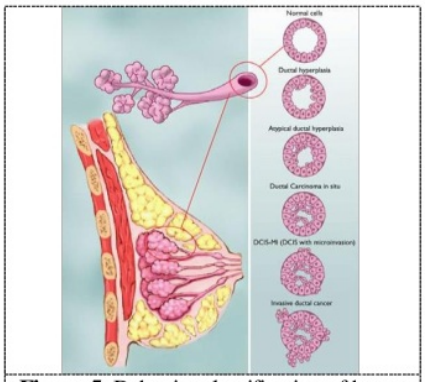

Figure 5. Behavior classification of breast cancer [14]

Table 5. Frequency distribution of breast cancer diagnosis by the behavior

Behavior	Frequency	Percent
Benign	72	25.5
Malignant	184	65.3
Metastase	26	9.2
Total	282	100.0

The division of breast cancer based on the grade is adjusted to the severity of breast cell modification disorders or differentiation. Grade 1 indicated well differentiated; grade 2 indicated intermediately differentiated, and grade 3 indicated poorly differentiated (Figure 6).

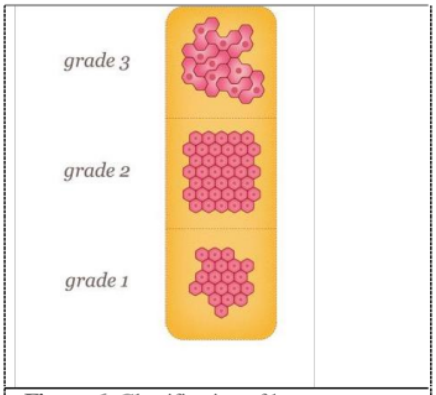

Figure 6. Classification of breast cancer according to grade [15]

Table 6. Frequency distribution of breast cancer diagnosis by grade of malignancy

Grade	Frequency	Percent
Well-differentiated	12	4.3
Intermediate differentiated	126	44.7
Poor differentiated	59	20.9
Undifferentiated	11	3.8
Not-determinated	14	5.0
Benign	60	21.3
Total	282	100.0

Table 6 shows that grade of breast cancer dominant in intermediate differentiated (44.7%) and only 20.9% data with the poorly differentiated grade. From the contingency table between types of diagnosis classification, chi-square value is calculated. The chi-square value explains the size of the proximity of each category variable. From the chi-square value, it will also be known the distance used to describe the points on the correspondence plot or perception map

Analysis of the relationship between the location of the tumor with each type, topography, morphology, behavior, and grade showed the results only location and type were significantly related

(X^2 = 12.324, p-value= 0.002). Correspondence analysis between variables found only the location relationship with topography which was not significantly related (X^2 = 24.557, p-value 0.138).

A significant relationship was found between topography and type of tumor (X^2 = 251.670, p-value 0.001); between types with morphology (X^2 = 180.833, p-value 0.001), between morphology with grades (X^2 = 350.316, p-value 0.001) and between grades with behavior (X^2 = 271.357, p-value 0.001).

Table 7. Results of correspondence analysis between types of diagnosis of breast cancer

Correspondence the diagnosis	Statistic test	
	Chi-square	p-value
Location with topography	24.557	0.138
Topography with Malignancy	251.670	0.001
Malignancy with morphology	180.833	0.001
Morphology with grade	350.316	0.001
Grade with behavior	271.357	0.001

The results of the analysis revealed a relationship between the location of cancer and the type of malignancy. There is a probability of a malignant or benign tumor-attacking the right or left breast. The correspondence analysis was carried out between the location of the tumor and topography, type of malignancy, morphology, grade, and behavior respectively. The results showed the significant correspondence between topography and type of malignancy, type of malignancy with morphology, morphology with grade, and grade with behavior.

4. Discussion

4.1. Correspondence analysis was carried out on six types of pathology diagnosis.

Diagnosis of this pathology is critical to direct the type of treatment that must be given to achieve recovery. It was found that the diagnosis of grades and behavior had high accuracy and correspondence values. It is realized that it is necessary to establish the classification of the diagnosis according to the severity of the pathology of breast cancer. In reality, this examination standard does not run entirely due to the problem of patient ability, doctor's skills and completeness of diagnostic examination facilities. This research can help if there is a problem of incomplete examination and low quality of the examination results. The significant relation between one type of examination and another is the decision key for surgery or treatment. That is, any level of given diagnosis could be made that should be related to the severity of the breast cancer and direct the treatment needed.

A close relationship between morphological, behavioral and grade diagnoses was found. Therefore, examination of the pathology package is a relatively accurate and useful examination because it has the value of treatment direction, choice of surgery, and prognosis of breast cancer [16].

42. Even so, of the many available examinations, there are still several other forms

Of diagnoses, such as staging histology, types of expression tumor proteins (ER / PR / HER), and BRCA genetic expressions. Thus the diagnostic examination of breast cancer can fully include clinical examination, laboratory, histopathology, cellular, genetic molecule (tumor gene expression signature) [9]. Breast cancer prevention solutions require early detection or early diagnosis and appropriate and prompt treatment. The urgency of early detection with the classification of breast cancer diagnosis allows early treatment as well so that there is an increase in healing while reducing the incidence of cancer, reducing complications, prolonging quality life expectancy, and reducing mortality. The

diagnosis of breast cancer begins with only a physical examination of a general practitioner that has progressed to the examination of genetic engineering by a cellular microbiology laboratory expert.

Meanwhile in the field of science and computing statistics developed various forms of modeling and analysis in predicting the possibility of occurrence of a phenomenon (dependent variable) based on one to several predictors. Various forms of test models are known, such as using correlation analysis, cluster analysis, regression analysis, especially logistic (binary, ordinal and multi-nominal), support vector machine, and artificial neural networks [8][17].

The combination of medical research and health science, statistical technology and statistical computing has brought several results which reveal some secrets about the causes, ways of diagnosis and the provision of more effective treatments. It is just that the fact shows that breast cancer is still a disease that increases its incidence, is still widely prevalent, the first rank of all cancers, and the main cause of death, both in Indonesia and on another part of the world [18][19].

43. The results of the study allow the accuracy of the diagnosis of breast cancer which states

The existence of a disease that requires classification according to the severity of breast cancer which will be the basis for adjusting the type of treatment or appropriate action that must be given. It is recommended to the researcher that to keep in mind the characteristics of the pathological features found in a microscopic examination of breast cancer specimens so that they can maintain the accuracy of diagnosis to get appropriate treatment or surgery.

5. Conclusion

Each type of diagnostic classification of breast cancer is entirely accurate and has a significant relationship with each other. It is recommended to perform a pathologic microscopic examination of breast cancer so that they can maintain the accuracy of diagnosis to get appropriate treatment or surgery.

References

[1] W. J. Gradishar *et al*., "Invasive breast cancer version 1.2016, NCCN clinical practice guidelines in oncology," *J. Natl. Compr. Cancer Netw*., vol. 14, no. 3, pp. 324–354, 2016.

[2] World Health Organization, "International Classification of Diseases for Oncology," Geneva, 2013.

[3] J. W. Berg, J. J. Decrosse, A. A. Fracchia, and J. Farrow, "Stromal sarcomas of the breast. A unified approach to connective tissue sarcomas other than cystosarcoma phyllodes," *Cancer*, vol. 15, no. 2, pp. 418–424, 1962.

[4] John Hopkins University, "Staging & Grade - Breast Cancer | Johns Hopkins Pathology," *JHU Medicine*, 2018. .

[5] F. Sadoughi, H. Lotfnezhad Afshar, A. Olfatbakhsh, and N. Mehrdad, "Application of Canonical Correlation Analysis for Detecting Risk Factors Leading to Recurrence of Breast Cancer.," *Iran. Red Crescent Med. J.*, vol. 18, no. 3, p. e23131, Mar. 2016.

[6] P. Yelland, "An Introduction to Correspondence Analysis," *Math. J.*, vol. 12, 2010.

[7] M. J. Greenacre, *Correspondence analysis in practice*.

[8] P. J. Lisboa and A. F. G. Taktak, "The use of artificial neural networks in decision support in cancer: A systematic review," *Neural Networks*, vol. 19, no. 4, pp. 408–415, May 2006.

[9] S. Ramaswamy *et al*., "Multiclass cancer diagnosis using tumor gene expression signatures.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 98, no. 26, pp. 15149–54, Dec. 2001.

[10] A. Zamani, A. M. Zamani, B. Amaliah, and A. Munif, "Implementasi Algoritma Genetika pada Struktur Backpropagation Neural Network untuk Klasifikasi Kanker Payudara," *J. Tek. ITS*, vol. 1, no. 1, pp. A222–A227, Sep. 2012.

[11] R. M. Levenson, E. A. Krupinski, V. M. Navarro, and E. A. Wasserman, "Pigeons (Columba livia) as trainable observers of pathology and radiology breast cancer images," *PLoS One*, vol.

10, no. 11, p. e0141357, 2015.
[12] K. M. Doll, A. Rademaker, and J. A. Sosa, "Practical Guide to Surgical Data Sets: Surveillance, Epidemiology, and End Results (SEER) Database," *JAMA Surg.*, 2018.
[13] B. Weigelt, F. C. Geyer, and J. S. Reis-Filho, "Histological types of breast cancer: how special are they?" *Mol. Oncol.*, vol. 4, no. 3, pp. 192–208, 2010.
[14] H. Kennecke *et al.*, "Metastatic behavior of breast cancer subtypes," *J. Clin. Oncol.*, vol. 28, no. 20, pp. 3271–3277, 2010.
[15] N. Houssami, P. Macaskill, M. L. Marinovich, and M. Morrow, "The association of surgical margins and local recurrence in women with early-stage invasive breast cancer treated with breast-conserving therapy: a meta-analysis," *Ann. Surg. Oncol.*, vol. 21, no. 3, pp. 717–730, 2014.
[16] N. Sourial *et al.*, "Correspondence analysis is a useful tool to uncover the relationships among categorical variables.," *J. Clin. Epidemiol.*, vol. 63, no. 6, pp. 638–46, Jun. 2010.
[17] F. Novianti, F. A. Novianti, and S. W. Purnami, "Analisis Diagnosis Pasien Kanker Payudara Menggunakan Regresi Logistik dan Support Vector Machine (SVM) Berdasarkan Hasil Mamografi Fourina Ayu Novianti dan Santi Wulan Purnami," *J. Sains dan Seni ITS*, vol. 1, no. 1, pp. D147–D152, Sep. 2012.
[18] P. L. Remington, R. C. Brownson, M. V. Wegner, and American Public Health Association, *Chronic disease epidemiology, prevention, and control*.
[19] H.-P. Sinn and H. Kreipe, "A Brief Overview of the WHO Classification of Breast Tumors, 4th Edition, Focusing on Issues and Updates from the 3rd Edition.," *Breast Care (Basel).*, vol. 8, no. 2, pp. 149–54, May 2013.

Correspondence Analysis of Breast Cancer Diagnosis Classification

ORIGINALITY REPORT

8%	6%	7%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Hamka Lodang, Nani Kurnia. "Distribution and Abundance of Plankton in The Downstream of Jeneberang River", Journal of Physics: Conference Series, 2019 Publication	2%
2	repository.tudelft.nl Internet Source	2%
3	mafiadoc.com Internet Source	1%
4	eprints.nottingham.ac.uk Internet Source	<1%
5	C. Basir, Y. Widyaningsih, D. Lestari. "Analysis of health in health centers area in Depok using correspondence analysis and scan statistic", AIP Publishing, 2017 Publication	<1%
6	Anastasia Pazaiti. "Basal phenotype breast cancer: implications for treatment and prognosis", Women s Health, 03/2011	<1%

	Publication	
7	"Colon Polyps and Colorectal Cancer", Springer Science and Business Media LLC, 2021 Publication	<1%
8	www.tandfonline.com Internet Source	<1%
9	hdl.handle.net Internet Source	<1%

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	On		

Correspondence Analysis of Breast Cancer Diagnosis Classification

GRADEMARK REPORT

FINAL GRADE

/0

GENERAL COMMENTS

Instructor

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7

PAGE 8

PAGE 9