Cox Proportional Hazard Survival Analysis to Inpatient Breast Cancer Cases

by Fatmah Afrianty Gobel

PAPER • **OPEN ACCESS**

Cox Proportional Hazard Survival Analysis to Inpatient Breast Cancer Cases

To cite this article: M. Nadjib Bustan *et al* 2018 *J. Phys.: Conf. Ser.* **1028** 012230

IOP Conf. Series: Journal of Physics: Conf. Series **1028** (2018) 012230 doi:10.1088/1742-6596/1028/1/012230

Cox Proportional Hazard Survival Analysis to Inpatient Breast Cancer Cases

M. Nadjib Bustan[1], Arman[2], M. Kasim Aidid[1], Fatmah Afrianty Gobel[2], and Syamsidar[2]

[1]Department of Statistics, Makassar State University, Makassar, 90222, Indonesia
[2]Post Graduate Program of Public Health, Universitas Muslim Indonesia, Makassar, 90232, Indonesia,

*mnbustan@unm.ac.id

Abstract: Breast cancer problems are characterized by increased incidence and trends, including death. The prognosis of survival rate in 5 years in stadium I is 90%, stadium II 65%, stadium III 15-20%, and stadium IV is less than 5%. To determine survival rate and cause of death, survival analysis has been done to find the relationship between survival rate and clinical stages, tumor location, metastasis, age and co-morbidity. A retrospective cohort study of inpatient breast cancer cases in Ibnu Sina UMI Hospital, Makassar, Indonesia 2013-2016 was selected as 108 cases out of 436 all inpatient cases. The survival analysis was performed using Kaplan-Meier and Cox Proportional Hazard methods. The results showed that the survival probability was 0.029, and the median survival was 20 months at intervals from 0 to 46 months. Significant related factors with survival rete were metastasis and co-morbidity. The Hazard Ratio of cancer stadium II dan III was 2,061 and hazard ratio of cancer stadium IVwas 1.571. Moreover, Stadium III survival rate was 73.1%, and stadium IV / metastases decreased to 6.5%, but it was increased to 42.6% when treated. Suggestion of the need for early detection of breast cancer, so women should come early to visit health facilities.

1. Introduction

Through Cancer, especially breast cancer is a very dangerous disease that spreads all over the world. [1]. Breast cancer gets a lot of doctors and statisticians' attention because it has a high mortality rate with an unstable tendency to death, with numerous and varied prognostic factors [1] [2] . The presence of breast cancer in Indonesia ranks first with a prevalence of 0.5 per 1,000 women. In addition to the high prevalence, breast cancer problems are characterized by high mortality, associated with late diagnosis or patients coming for treatment when it goes to advanced stage [3] [4] . Classification of breast cancer stage according to AJCC [5] has a relationship with survival prognosis. This situation is indicated by decreasing survival rate from Stage I 90%, Stage II 65%, Stage II 20% and Stage IV 5% [6].To determine why the survival rate of breast cancer cases is low, it is necessary

IOP Conf. Series: Journal of Physics: Conf. Series **1028** (2018) 012230 doi:10.1088/1742-6596/1028/1/012230

to identify the prognostic factors survival rate [7]. The major prognostic factors of breast cancer are associated with several possible approaches to diagnosis, including clinic stages [8], and histo-pathological grade [9] [10]. Survival rate depends also on the type of diagnosis based on the existence of metastasis [11-13]; radiologic figure of mammography [14]; and genetic expression [15]. In addition to the clinical stage stage of breast cancer or the type of diagnosis, there are several other prognostic factors associated with survival rate, namely: age [16], race [15] , occupation [17], co-morbidity: diabetes [18]; anemia [7] [19], family history [7], tumor size [20] [21], location of lymph node metastasis [11] [21]. Survival rate is also related to type of surgery and treatment, including in chemotherapy [22], breast surgery [23], hormonal [7] and radiotherapy [24].

Therefore, this study will analyze the prognostic factors associated with survival rate of breast cancer patients. Based on breast cancer condition, especially in predicting the important prognostic factors, statistician could used survival analysis, which have several models. One of most commonly used model is Cox proportional hazard with Kaplan-Meier and Log-Ranks test [24] [25] [26-28].

2. Methods and Materials

2.1 Study Design

Learning is A retrospective cohort study of inpatient breast cancer cases in IbnuSina UMI Hospital, Makassar, Indonesia 2013-2016 was selected as 108 cases out of 436 all inpatient cases. The survival analysis was performed using Kaplan-Meier and Cox Proportional Hazard methods [29] [30]. Kaplan-Meier model was used to determine survival probality of breast cancer patients. Then, Log-Rank test was used to determine the significance difference between survival expression of the patient [26]. Cox proportional hazard was used to determinethe difference ratio of prognostic factors which were including age, clinical stages, tumor size, location of lymph node metastasis, and co-morbidity [31].

2.2 The Survival Function

Individual opportunities to survive for time x are expressed by $S(x) = P(\mathbf{X} > x)$. Let **X** be the continuous random variables, then the survival function is the complement of the cumulative distribution function $S(x) = 1 - F(X)$ where $F(X) = P(\mathbf{X} \leq x)$. The survival function is the integral of the probability density function $f(x)$:

$$S(x) = P(X > x) = \int_x^\infty f(t)\,dt \qquad (1)$$

$$f(x) = -\frac{dS(x)}{dx} \qquad (2)$$

Then if **X** is the discrete random variables, and can be obtained x_j, j=1,2,3,… with the probability mass function (p.m.f) $p(x_j) = P(\mathbf{X} = x_j)$, j=1,2,3,… where $x_1 < x_2 < x_3$ … then the survival function for the discrete variables **X** is given by:

$$S(x) = P(X > x) = \sum_{x_j > x} p(x_j) \qquad (3)$$

2.3 Hazard Functions

The hazard function of the hold time **X** is denoted by h(x) and defined as individual probability fails in the time interval $(x, x+\Delta x)$ that the individual has lived for time x, the hazard function is expressed as:

$$h(x) = \lim_{\Delta x \to 0}\left[\frac{P(x < X < x + \Delta x \mid X > x)}{\Delta x}\right] \qquad (4)$$

The relationship between the hazard function and the survival function is expressed by:

$$h(x) = \frac{f(x)}{S(x)} \tag{5}$$

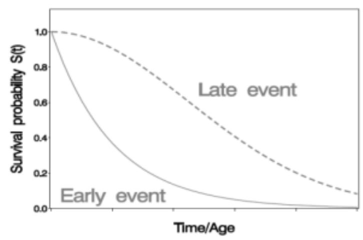

Figure 1. Survival Opportunities For Early Event and Late Event Based On Time.

2.4 Kaplan Meier's Survival Curve and Log Rank Test

Estimated Kaplan Meier survival function. expressed by:

$$\hat{S}(x_{(j)}) = \hat{S}(x_{(j-1)})\hat{P}\left[X > x_{(j)} \mid X \geq x_{(j)}\right] \tag{6}$$

A further log rank test is used to compare Kaplan Meier's survival curves formed by the following hypothesis:

H_0: There is no difference between the survival curves

H_1: At least one difference between the survival curves:

$$\text{Log Rank Statistic } = \frac{(O_i - E_i)^2}{Var(O_i - E_i)} \tag{7}$$

with

$$O_i - E_i = \sum_{j=1}^{n}(m_{ij} - e_{ij}) \tag{8}$$

m_{ij} denotes the number of individuals who experience the event at time x_j, and eij is the value of hope. The null hypothesis will be rejected if log rank statistics $\geq \chi^2_{\alpha,df}$ with degrees of freedom (df) = 1 or p-value <α.

2.5 Cox Proportional hazard model

The relationship between the hazard rate and the covariate set can be expressed using the model:

$$\ln[h(t)] = \ln[h_0(t)] + \sum_{i=1}^{n} x_i\beta_i \tag{9}$$

$$\text{or} \quad h(t) = h_0(t)e^{\sum_{i=1}^{n} x_i\beta_i} \tag{10}$$

Where $x_1, x_2, \ldots, x_n$ are covariates. $\beta_1, \beta_2, \ldots, \beta_n$ are the regression coefficients to be estimated. t is time and $h_0(t)$ is the baseline hazard rate when all covariates are zero.

2.6 Concept Framework and ResearchVariables

Potential independent variables affecting patient survival are tumor size, clinical stage, history of metastasis, co-morbidities, and age. Survival; the probability of breast cancer survivors to observe 3 (three) years (2013 - 2016): 1 means death (event) and 0 means life (censorship). Tumor Size; breast tumor diameter T1 (<20 cm), T2 (2-5 cm), T3 (> 5 cm). Clinical Stage; degree or degree in breast cancer patients; consisting of 1 (Stage IV), 2 (stage IIIA, IIIB, IIIC), 3 (stage I, IIA, IIB), and 4 (stage 0). History of metastasis; cancer cells that have reached to other organs with objective criteria that are present and nonexistent.Complicated disease; the presence of other diseases of breast cancer.Age; age of breast cancer patients at the beginning of treatment where 1 (60 years), 2 (41-60 years), and 3 (<40 years).

3. Results and Discussion

3.1. Kaplan Meier Analysis

The probability of survival of breast cancer patients in Ibnu Sina Hospital Makassar in 2013-2016 was indicated by unit months (0-46 months) as in table 1.

Table 1. Probability of Survival According to Tumor Size, Clinical Stage, Metastasis, Comobidity, and Age

Survival Rate		Tumor Size		Stadium			Metastases		Comorbidity		Age (year)		
Time ((months)	Cumulative Proporsional survival	2-5 cm	>5 cm	(I, IIA, IIB)	(IIIA, IIIB, IIIC)	IV	Yes	No	Yes	No	< 40	41 - 60	> 60
0	0,861	0	0,853	0	0,835	0,714	0,674	0	0,733	0,910	0,864	0,870	0,78
5	0,796	0	0,784	0	0,759	0	0,565	0,968	0,667	0,846	0,818	0,818	0
10	0,741	0	0,725	0	0,696	0	0,478	0	0,567	0	0,682	0	0
15	0,611	0	0,588	0	0,532	0	0,348	0,806	0,900	0,692	0	0,623	0
20	0,491	0	0,471	0,818	0,418	0	0,261	0,661	0,233	0,590	0	0,494	0
25	0,370	0	0,343	0,727	0,278	0	0,217	0,484	0	0,436	0,409	0,403	0
30	0,293	0,625	0,272	0,630	0,199	0	0,152	0,399	0,20	0,357	0	0,335	0
35	0,184	0	0,164	0,350	0	0	0	0,241	0	0,222	0,212	0	0
41	0,086	0	0,068	0,210	0,050	0	0	0,124	0	0,113	0	0,111	0
46	0,029	0	0,023	0	0	0	0	0	0	0,380	0	0,370	0
Median	20	24,875	19,897	33,00	17,00	11,00	9,00	25,00	11,00	24,00	24,00	20,00	16,00

Probability of survival reached 46 months by 0.029 (2.9%). This means that if there are 100 people sufferers, then that can survive up to 46 months only 3 people only; and decreased drastically (Figure 2).

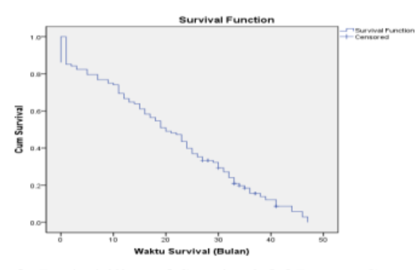

Figure 2. Probability of Survival Of Breast Cancer Patients

IOP Conf. Series: Journal of Physics: Conf. Series **1028** (2018) 012230 doi:10.1088/1742-6596/1028/1/012230

Figure 3 showed that there is a drastic decrease in survival chances of breast cancer patients who have a history of metastasis compared with patients without a history of metastasis. Meanwhile, patients with breast cancer who do not have co-morbidities are higher in survival than patients with coexisting illness (Figure 4). One of the complaints of breast cancer patients is shortness of breath caused by metastasis to the lungs causing other effects such as pleural effusion.

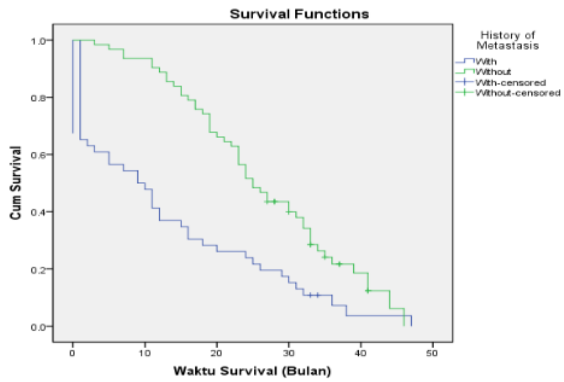

Figure 3. Probability of Survival of Breast Cancer Patients According to History Of Metastasis

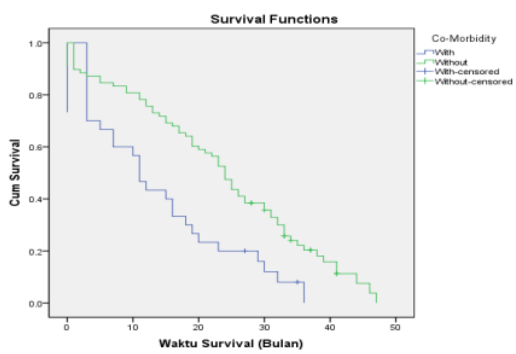

Figure 4. Probability of Survival of Breast Cancer Patients According to Co-Morbidities.

3.2. Rank Log Test

To determine whether there are differences between the survival curve then used the log rank test.

Table 2. Rank Log Test Factors Affecting Survival of Breast

Variable	Log rank (Chi-Square)	df	P-value
Tumor size	5.045	1	0.025
Clinical Stadium	11.259	2	0.004
Metastases	15.860	1	0.000
Co-morbidity	11.616	1	0.001
Age	7.763	2	0.021

Based on log rank test in table 2, it can be seen that the survival of breast cancer patients based on tumor size variables, clinical stage, history of metastasis, comorbidities and age were statistically significantly different with p-value <0.05.

3.3 Cox Proportional Hazard Model

For all tumor-size factors (x1), clinical stage (x2), history of metastasis (x3), comorbidities (x4) and age (x5) allegedly affecting survival of breast cancer patients are generally modeled as:

$$h(t) = h_0(t)e^{(\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5)}$$

Based on table 4 in step 1 obtained model:

$$h(t) = h_0(t)e^{(0.87x_1 + 0.38x_{21} + 0.4x_{22} + 0.45x_3 + 0.50x_4 + 0.22x_5)}$$

In step 1, the partial test shows that only PP factors are statistically significant (P-value <5%). The backward stepwise method is used to extract the least influencing factors so that the final model is obtained in step 4:

The βregression coefficients of the obtained models are all positive (β>0) with the value of expβ>0, meaning that all factors included in the model influence the event speed (death). That is, the risk of failure (death) of breast cancer patients with metastasis is 1.94 times greater than that not metastases. The risk of death of breast cancer patients with comorbidities is 1.7 times greater than those that do not have co-morbidities.

Table 3. Partial Test With Backward Stepwise Method

Variable		β	Wald	p-value	Exp (β)
Step 1	x_1	0.868	1.316	0.251	2.383
	x_2		1.999	0.368	
	x_2 (1)	0.380	0.525	0.469	1.462
	x_2 (2)	0.452	1.990	0.158	1.571
	x_3	0.447	3.450	0.063	1.564
	x_4	0.498	4.147	0.042	1.646
	x_5	0.217	0.911	0.340	1.242
Step 2	x_1	1.165	2.605	0.107	3.207
	x_3	0.588	6.992	0.008	1.800
	x_4	0.501	4.174	0.041	1.650
	x_5	0.213	0.866	0.352	1.238
Step 3	x_1	1.157	2.568	0.109	3.180
	x_3	0.586	6.917	0.009	1.797
	x_4	0.522	4.543	0.033	1.686
Step 4	x_3	0.663	8.941	0.003	1.941
	x_4	0.535	4.725	0.030	1.707

The result of Cox PH analysis showed that the most significant and significant prognostic factor to the probability of death was the presence of metastases and co-morbidities.

4. Conclusion

The main factor causing low survival time was because the patient comes for treatment already in an advanced stage even accompanied by comorbidities (such as diabetes, anemia and hypertension). It is recommended that health workers conduct promotions to motivate women at risk for early self-examination if they know of any signs of a tumor in the breast.

References

[1] Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A 2015 Global cancer statistics, 2012. *CA Cancer J Clin* **65(2)**:87–108.

[2] Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al 2015 Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer **136(5)** E359–86.

[3] Ayu G, Dyanti R, Luh N, Suariyani P 2016 Delaying Factors In Breast Cancer Patients Taking Early Examination Into Health Services. KEMAS 11(2).

[4] Bustan MN. Manajemen Pengendalian Penyakit Tidak Menular [Internet]. Rineka Cipta; 2015

[5] American Cancer Society. Stages of Breast Cancer [Internet]. [cited 2017 Sep 28].

[6] Sinaga ES, Ahmad RA, Hutajulu SH 2017 Berita kedokteran masyarakat. Vol. **33**, Fakultas Kedokteran, Universitas Gadjah Mada.

[7] Álvarez-Bañuelos MT, Rosado-Alcocer LM, Morales-Romero J, Román-Álvarez LS, García REG-, Carvajal-Moreno M 2016 Prognostic Factors Associated with Survival in Women with Breast Cancer from Veracruz, Mexico. *J Cancer Sci Ther* Apr 23 **8**(4).

[8] Saadatmand S, Bretveld R, Siesling S, Tilanus-Linthorst MMA 2015 Influence of tumour stage at breast cancer detection on survival in modern times: population based study in 173 797 patients. BMJ 351.

[9] Klauber-DeMore N. Tumor Biology of Breast Cancer in Young Women. DeMore N, editor. Breast Dis. 2006 Jun 23 **23(1)**:9–15.

[10] Azim HA, Partridge AH 2004 Biology of breast cancer in young women. Breast Cancer **16(4)**:427.

[11] Kuru B, Camlibel M, Dinc S, Gulcelik MA, Gonullu D, Alagol H 2008 Prognostic factors for survival in breast cancer patients who developed distant metastasis subsequent to definitive surgery. Singapore Med J [Internet] **49(11)** 904–11. 7

[12] Elston CW, Ellis IO, Pinder SE. Pathological prognostic factors in breast cancer. Crit Rev Oncol . 1999;31:209–23.

[13] Chang J, Clark GM, Allred DC, Mohsin S and Chamness G 2003 Elledge RM. Survival of patients with metastatic breast carcinoma. Cancer **97(3)**:545–53.

[14] Lee SC, Jain PA, Jethwa SC, Tripathy D, Yamashita MW 2014 Radiologist’s Role in Breast Cancer Staging: Providing Key Information for Clinicians. RadioGraphics **34(2)**:330–42.

[15] Carey LA, Perou CM, Livasy CA, Dressler LG, Cowan D, Conway K, et al 2006 Race, Breast Cancer Subtypes, and Survival in the Carolina Breast Cancer Study. JAMA **295(21)** 2492.

[16] Fredholm H, Eaker S, Frisell J, Holmberg L, Fredriksson I and Lindman H 2009 Breast Cancer in Young Women: Poor Survival Despite Intensive Treatment. Aziz SA, editor. PLoS One **4(11)**:e7695.

[17] Karjalainen S and Pukkala E. Social class as a prognostic factor in breast cancer survival. Cancer [Internet]. 1990 Aug 15 [cited 2017 Sep 26];**66(4)**:819–26.

[18] Lipscombe LL, Goodwin PJ, Zinman B, McLaughlin JR and Hux JE 2008 The impact of diabetes on survival following breast cancer. Breast Cancer Res Treat **109(2)**:389–95.

[19] Nechuta S, Lu W, Zheng Y, Cai H, Bao P-P, Gu K, et al. Comorbidities and breast cancer survival: a report from the Shanghai Breast Cancer Survival Study. Breast Cancer Res Treat . 2013 May 19;**139(1)** 227–35.

[20] Abner AL, Collins L, Peiro G, Recht A, Come S, Shulman LN, et al 1998 Correlation of tumor size and axillary lymph node involvement with prognosis in patients with T1 breast carcinoma. Cancer **83(12)**:2502–8.

[21] Carter CL, Allen C, Henson DE 1989 Relation of tumor size, lymph node status, and survival in 24,740 breast cancer cases. *Cancer* **63(1)** 181–7.

[22] Gagliato D de M, Gonzalez-Angulo AM, Lei X, Theriault RL, Giordano SH, Valero V, et al

2014 Clinical impact of delaying initiation of adjuvant chemotherapy in patients with breast cancer. *J Clin Onco*l **32(8)**:735–44.

[23] Kim KJ, Huh SJ, Yang J-H, Park W, Nam SJ, Kim JH, et al. Treatment Results and Prognostic Factors of Early Breast Cancer Treated with a Breast Conserving Operation and Radiotherapy.

[24] Abadi A, Yavari P, Dehghani-Arani M, Alavi-Majd H, Ghasemi E, Amanpour F, et al 2014 Cox models survival analysis based on breast cancer treatments. Iran J cancer Prev **7(3)**:124–9.

[25] Pham M 2014 Survival Analysis - Breast Cancer. *Undergrad J Math Model One Two.2014 Jan* **1**;6(1).

[26] Usman M, Bala S, Dikko H and Gulumbe S 2014 An application of kaplan-meier survival analysis using breast cancer data. *Sub-Saharan African J Med* **1(3)**:132.

[27] Mangasarian OL, Street WN, Wolberg WH. Breast Cancer Diagnosis and Prognosis via Linear Programming. 1994. Spring/1994/SS-94-01/SS94-01-019.

[28] Collett D. Modelling survival data in medical research. 532 p.

[29] Lee ET, Wang JW. Statistical methods for survival data analysis.

[30] Kleinbaum DG 1998 Survival Analysis, a Self-Learning Text. *Biometrical J* Apr;**40(1)**:107–8.

[31] Soerjomataram I, Louwman MWJ, Ribot JG, Roukema JA and Coebergh JWW 2008 An overview of prognostic factors for long-term survivors of breast cancer. Breast Cancer Res Treat **107(3)**:309–30.

Cox Proportional Hazard Survival Analysis to Inpatient Breast Cancer Cases

ORIGINALITY REPORT

15%	11%	6%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

MATCH ALL SOURCES (ONLY SELECTED SOURCE PRINTED)

1%

★ docplayer.net

Internet Source

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		