Prognostic Cut Point for Breast Cancer Age of Diagnosis

Ebrahim Hajizadeh,1 Mahbube Abdollahi,1 Ahmad Reza Baghestani,2 and Shahpar Haghighat3,*

1Department of Biostatistics, Faculty of Medical Sciences, Tarbiat Modares University, Tehran, IR Iran
2Department of Biostatistics, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, IR Iran
3Quality of Life Department, Breast Cancer Research Center, Motamed Cancer Institute, ACECR, Tehran, IR Iran

Abstract

Background: In oncology studies, categorizing a quantitative prognostic variable or determining cut point is aimed at categorizing individuals into homogenous groups. Such categorizations are useful for treatment recommendations and clinical trial design.

Objectives: This article aims at determining the cut point for breast cancer diagnosis age by factors affecting the patients’ survival, using cure model.

Methods: In this longitudinal study, a total 559 patients with breast cancer referring to breast cancer research center, Tehran, Iran, from 1986 to 2006 entered the study. Patients were followed until 2013. The last status of patients was recorded, using telephone conversation. Then, the cut point of breast cancer diagnosis age was determined, using change point cure model with survival-related covariates and R v.2.15.0 software.

Results: In the present study, the mean age of diagnosis was reported 46.31 ± 11.17. Median time of follow-up was 68.36 months with the range of 0.89 to 324. The results showed that age cut point was 49.45 ± 0.64. In young group, one unit increase in tumor size led to 57% reduction in the chance of cure. In old group, the chance of recovery declined by 51%. In old group, the chance of cure among those with lymph node declined by 56% compared to those without lymph node. In the young group, this variable was not significant. Level of education, type of surgery, and estrogen receptor had no significant relationship with cure in none of the age groups at 5% error level.

Conclusions: The results showed that the effect of age in breast cancer prognosis is adjusted at 50-year cut point, leading to two relatively homogeneous groups. This cut point is effective in assessing predictive and prognostic factors in breast cancer. The difference between effect of tumor size and effect of lymph node involvement in different age groups can be helpful in determining more appropriate therapeutic strategies.

Keywords: Breast Neoplasms, Diagnosis, Patients

1. Background

Breast cancer is a multi-stage cancer, which is curable if diagnosed in early stages (1). In the 21st century, it seems that we are witnessing a progressive decline in deaths from the disease, especially in developed countries (2). Breast cancer is the second most common cancer worldwide and the fourth cause of death from cancer. Breast cancer is the most common cancer among women; 1.67 million new cases of cancer have been diagnosed in 2012 (3). In Iran, breast cancer is the most prevalent cancer (23% of all cancers) among women (4). Progress in the field of cancer is one of the main goals of health care programs (5). In oncology studies, classifying a quantitative prognostic variable or determining cut point is requested to categorize individuals into low-risk and high-risk classes. Therefore, homogenous groups of patients are created. Such categorizations are useful for treatment recommendations, clinical trial design, and treatment improvement planning (6, 7). There are multiple methods to categorize a quantitative variable or determine cut point such as using the sample quantiles (median), the optimal change point, change-point method, etc. (7, 8). Multiple studies have been conducted to determine the cut point, using change-point models. Tashima et al. (2015) conducted a study to determine ki-67 cut point in patients with breast cancer. The article aims at introducing ki-67 indicator as a more accurate diagnostic factor. In this study, comparison of P values and hazard ratios in univariate and multivariate cox proportional hazards (Cox PH) model and were employed to determine the optimal cut point (9). Syed K Moshin et al. (2004) conducted a study in patients with breast cancer to develop and validate immune histochemical assay method. In this study, minimum P value method was used to determine the optimal cut point for positive progesterone receptor (10). Lopez et al. (2010) conducted a study to up-
date changes in colon cancer mortality and incidence. In this study, transition change-point model was used. They concluded that the incidence of colorectal cancer has increased markedly compared to the past. The increased incidence and mortality rates were in opposite direction (11). Since breast cancer is one of preventable diseases with screening (1), knowing change point, biological, and prognostic characteristics are highly regarded in a screening planning policy making (5). Therefore, in this study, change-point cure model is employed to estimate change point (cut point) in diagnosis age of breast cancer.

This model was introduced by Othus et al. in order to estimate the change point in a quantitative variable in long-term survival data (12). In this model, change point in breast cancer age is obtained by dividing likelihood function of a non-mixture cure model into two sections (prior to and after change point). Cure model is used because using standard models such as semi-parametric Cox PH models are not appropriate in survival studies if there are long-term or cured individuals. These models assume that all individuals would experience the event (13, 14).

Therefore, cure models are used in such cases. Some factors have been introduced as survival prognostic factors in breast cancer, including higher stage of disease, number of lymph nodes involved in cancer, pathology, socioeconomic status, type of treatment, etc. (15). These factors need to be taken into account in determining a prognostic model. Disease free survival (DFS) time is one of the most common criteria to evaluate the impact of various therapeutic agents in patients. DFS time is defined as duration from the onset of sickness to event (death or recurrence) (16).

This article aims at estimating the change point of cancer diagnosis age, using change point cure model with some predictive factors including level of education, estrogen receptor, lymph node involvement, etc. which are effective factors in survival. It also studies the concurrent effect of predictive factors prior to and after change point on cure rate. Determining cut point and distribution and relationship between variables and event prior to and after this point can be a valuable solution for clinicians and health policy makers to develop a cost-effective treatment guideline in patients diagnosed with breast cancer.

2. Methods

In this longitudinal study, 559 patients with breast cancer referring to breast cancer research center, Tehran, Iran from 1986 to 2006 entered the study. The patients were re-followed in 2013. Demographic information and clinical documentation were completed. Their survival was recorded through telephone conversation. The date of cancer diagnosis was considered the onset of disease. DFS time was calculated based on the telephone conversation follow-up. Change point cure model was employed to determine a cut point in age. The method used in this article is according to a new change point model introduced by Othus et al. (12). In that study, introduced likelihood function was tested by a simulation study and goodness of fit of the model was tested. Then, the model was used to determine a cut point in diagnosis age of prostate cancer (12). In exponential cure model, cut point of breast cancer diagnosis age was estimated, using some predictive factors, including level of education, lymph node involvement, tumor size, status of estrogen and progesterone receptors, type of surgery, and chemotherapy and radiotherapy, which are effective factors in survival.

2.1. Mixture Cure Model with Covariates

In mixture cure model, it is assumed that the population consists of two groups (cured and non-cured) and survival function of mixture cure model is defined based on this assumption. Survival function is as follows:

\[ s(t) = p + (1 - p) S_u(t) \] (1)

\[ S_U(t) \] is survival function for non-cured patients. For this model, all individuals are alive at \( t = 0 \) (\( S(0) = 1 \)). At infinite, a proportion of individuals is cured (\( S(\infty) = p \)). In other words, some do not experience event (here recurrence of cancer). The presence of individuals with long-term survival makes the Kaplan-Meier curves not to reach zero at the end.

It remains horizontally constant. Prior to using cure model, two pre-assumptions need to be studied (presence of cured people and sufficient length of follow up). In case of using covariates \( X \), logistic link function is used as follows (14, 17, 18).

\[ p = \frac{\exp(BX)}{1 + \exp(BX)} \] (2)

Where \( \beta \) shows the effect of covariates on chance of cure. So, \( e^{\beta} \) shows odds ratio for covariate \( X \).

2.2. Change Point Cure Model with Covariates

Mixture cure model with a change point at a quantitative variable: In this model, likelihood function of mixture cure model is divided into two sections (prior to and after change point) and change point of \( X \) quantitative variable is estimated.

In a mixture cure model with change point at a covariate variable, density function is written as follows:

\[ f = \left( f_1(t)^{d} S_1(t)^{1-d} \right)^{I(X<\tau)} \left( f_2(t)^{d} S_2(t)^{1-d} \right)^{I(X>\tau)} \] (3)

Hajizadeh E et al.
Where \( f(t) \) is density function of time variable until the event before and after change point parameter, \( \tau \) is change point parameter for X quantitative variable, \( S(t) \) is survival function of time variable until the event before and after change point, \( d \) is status variable, censor for each individual so that in case of event, \( d = 1 \), if censored, then, \( d = 0 \). Above function is not smooth.

Therefore, final likelihood function is as follows after smoothing (12).

Maximum likelihood method and numerical search methods are employed to estimate the change-point model, including \( \beta \) coefficient and \( \tau \) change-point parameter.

2.3. Data Analysis

SPSS v.16 was employed to draw KM curve. R and Stats4 package were employed to estimate the change-point model parameters. First, the appropriateness of cure model was studied, using likelihood ratio test. Then, likelihood function was written for change-point cure model with covariates. Finally, maximum likelihood estimation was carried out, using numerical search methods.

This study was approved by ethics committee of Tarbiat Modares University.

3. Results

A total of 559 women entered the study. The mean diagnosis age of patients, who were analyzed, was reported 46.31 ± 11.17 years. Patients’ age varied between 24 and 83 years. Mean and median time of follow-up was 81.43 ± 55.72 and 68.36 months, receptively. The range of changes was 0.89 - 324. Totally, 179 (32%) deaths occurred and the rest were considered right censored.

Table 1 shows the demographic characteristics of patients who were analyzed. As it can be seen, almost 51% (287 patients) had level of education lower than diploma; 64% (358 patients) experienced lymph node involvement. The mean size of tumor was 2.17 ± 1.05 cm; 83.4% of patients (316 patients) underwent modified radical mastectomy (MRM) and the rest of patients experienced breast preservation (BP). As much as 56% (316 patients) had positive estrogen receptor. The mean tumor size, who underwent MRM, was 2.26 ± 1.09 cm. It was reported 1.71 ± 0.73 in women with breast preservation.

Table 2 shows the change point model fitting results. Confidence interval for OR was calculated, using the following formula. The results indicate that breast cancer change-point age is 49.45 (± 0.64) years. In this study, lower than diploma, lack of lymph node involvement, Breast Preservation surgery, and negative estrogen receptor were considered as reference group. The results of cure change-point model indicate that the chance of cure was 50% less in younger group with lymph node than those without lymph node and this variable was not significant. However, in older group, the chance of cure declined by 61% and it was significant. In younger group, one unit of tumor size increase led to 57% decline in chance of cure. In older group, one unit in tumor size increase led to 51% decline in chance of cure and tumor size variable was significant in both age groups. Level of education, type of surgery, and estrogen receptor variables were not significant in none of age groups. In younger group, the chance of cure was 8% lower in diploma and higher level of education group than that of lower-than-diploma group. In older group, however, the chance of cure was 56% higher. In younger group, the chance of cure in MRM surgery was 1.3 times as much as BP surgery. In older group, the chance of cure was twice. In younger group, the chance of cure was 50% higher in positive estrogen receptor than negative estrogen recep-
In different studies, some factors are considered prognostic ones such as younger age, comorbidity, larger tumor size, negative estrogen receptor, higher number of lymph node involvement that in hazard estimation models, a significant relationship is found with relapse (15, 25). In this study, lymph node involvement in older-than-50 age group displayed a significant relationship with survival. But, this relationship was not significant in younger ages at error of 5%. In a study conducted by Wildiers et al. the probability of lymph node involvement declines until 70 years old. This probability, however, rises after 70, which is consistent with this study (26). Using different cut points and lack of relationship between survival prognostic variables and adjustment of effect of age can be associated with lack of evidence. Mamunas raised the same criticism regarding this probability (25). In this study, we obtained change point for breast cancer diagnosis age, using change point mixture cure model.

Considering the nature of long-term survival data or presence of cured patients in the study is one of the advantages of this model. We were also able to estimate and compare the covariate-related parameters affecting survival in both age categories. Determining change point is carried

\[ L = \prod_{i=1}^{n} \left( (1 - p_1) \times f_1(t_i) \times \delta_i (p_1 + (1 - p_1) \times S_1(t_i))^{1-\delta_i} \right) \frac{\exp \left( -\frac{X - \tau}{n(-\frac{1}{2})} \right)}{1 + \exp \left( -\frac{X - \tau}{n(-\frac{1}{2})} \right)} \left( (1 - p_2) \times f_2(t_i) \times \delta_i (p_2 + (1 - p_2) \times S_2(t_i))^{1-\delta_i} \right) \]

\[ e^{\hat{\beta} \pm s_b \times Z_{p \cdot 075}} \]

**4. Discussion**

This study aimed at determining the cut point of breast cancer diagnosis age with covariates. The cut point was obtained 49.45. The results also showed that tumor size increase in young group and lymph node involvement in old group declined the chance of recovery. Determining cut point in a quantitative variable or categorizing a quantitative variable is widely used in different sciences. Contal et al. (1999) employed two different methods in semiparametric Cox survival model to determine the cut point of breast cancer diagnosis age. They obtained the 41 and 37 year-old cut point by these two methods, respectively which is inconsistent with this study (19). Different cut points have been employed in studies concerning breast cancer. For example, 50 cut point (20-22) were employed in some studies conducted in the USA and Spain, which is line with the present study. Some studies used different cut points for age (15, 23, 24). The cut points used in these studies are not consistent with this study. This difference might be associated with different statistical population and different methods to obtain cut point.

In Li Christopher et al.’s study that reviewed multiple studies concerning breast cancer, the cut point of 50 years was introduced as prognostic factor. The survival probability of patients older than 50 years is 70% to 85%, while it is 55% to 80% for patients younger than 50 years. In this study, tumor size and lymph node involvement have been suggested as prognostic survival factors in younger-than-50 age groups; however, in addition to tumor size, lymph node involvement is considered as a lower survival prognostic factor in older-than-50 age group. Individuals who receive positive estrogen and progesterone have higher chance of survival due to effectiveness of wide variety of hormonal treatments compared to those who are negative receptors (15).

In different studies, some factors are considered prognostic survival factors such as younger age, comorbidity, larger tumor size, negative estrogen receptor, higher number of lymph node involvement that in hazard estimation models, a significant relationship is found with relapse (15, 25). In this study, lymph node involvement in older-than-50 age group displayed a significant relationship with survival. But, this relationship was not significant in younger ages at error of 5%. In a study conducted by Wildiers et al. the probability of lymph node involvement declines until 70 years old. This probability, however, rises after 70, which is consistent with this study (26). Using different cut points and lack of relationship between survival prognostic variables and adjustment of effect of age can be associated with lack of evidence. Mamunas raised the same criticism regarding this probability (25). In this study, we obtained change point for breast cancer diagnosis age, using change point mixture cure model.

Considering the nature of long-term survival data or presence of cured patients in the study is one of the advantages of this model. We were also able to estimate and compare the covariate-related parameters affecting survival in both age categories. Determining change point is carried
Table 2. Parameter Estimations and Confidence Intervals in Change Point Exponential Cure Model

| Variable | OR   | CI (95%)         |
|----------|------|------------------|
| Younger group (age ≤ 50) |      |                  |
| Hazard rate | 0.008 | (0.005, 0.011)   |
| Education | 0.413 | (0.413, 1.952)   |
| Diploma and higher | 0.917 |                  |
| Lower than diploma | Reference |          |
| Lymph node involvement | Yes | 0.520 (0.236, 1.145) |
| No | Reference |                  |
| Tumor size | 0.430 | (0.235, 0.787)   |
| Surgery type | (0.453, 3.803) |      |
| Modified Radical Mastectomy | 1.313 |                  |
| Breast preservation | Reference |          |
| Estrogen receptor | Yes | 0.748 (0.329, 1.701) |
| No | Reference |                  |
| Older group (age > 50) |      |                  |
| Hazard rate | 0.019 | (0.012, 0.027)   |
| Education | 0.019 | (0.012, 0.027)   |
| Diploma and higher | 1.560 | (0.632, 3.853)   |
| Lower than diploma | Reference |          |
| Lymph node involvement | Yes | 0.397 (0.168, 0.941) |
| No | Reference |                  |
| Tumor size | 0.430 | (0.235, 0.787)   |
| Surgery type | (0.453, 3.803) |      |
| Modified Radical Mastectomy | 2.093 |                  |
| Breast preservation | Reference |          |
| Estrogen receptor | Yes | 0.748 (0.329, 1.701) |
| No | Reference |                  |
| Change point at age of cancer diagnosis | 49.445 (48.131, 50.759) |

The results of the study show that the effect of age is adjusted at 50 years old in prognosis of patients with breast cancer, creating two relatively homogeneous groups. This cut point can be useful to evaluate various predictive and prognostic factors of breast cancer. The difference in the effect between tumor size and lymph node involvement in different age groups can be helpful in determining more appropriate therapeutic strategies.

Acknowledgments

This research is a part of Ph.D. dissertation of Tarbiat Modares University, which has financially been supported by research deputy of TMU. A debt of gratitude is owed to breast cancer research center, ACECR, Tehran, Iran to provide the data for this paper.

Footnotes

Authors’ Contribution: Mahbube Abdollahi designed the study, analyzed the data, and wrote the paper. Ebrahim Hajizadeh contributed to literature review, study design, and analysis. Shahpar Haghighat contributed to literature review and writing-up process. All authors read and approved the final manuscript.

Conflict of Interest: No conflict of interest.

References

1. Parker WH. Uterine fibroids. Berek and Novak’s Gynecology. 2012.
2. Forouzanfar MH, Foreman KJ, Delossantos AM, Lozano R, Lopez AD, Murray CJ, et al. Breast and cervical cancer in 187 countries between 1980 and 2010: a systematic analysis. Lancet. 2011;378(9801):1461–84. doi: 10.1016/S0140-6736(11)61351-2. [PubMed: 21924486].
3. Globocan I, Ferlay J, Bray F, Pisani P, Parkin D. Cancer incidence, mortality and prevalence worldwide. IARCPress. GLOBOCAN; 2000.
4. Khosravi A, Aghamohamadi S, Kazemi E, Pour Malek F, Shariati M. Mortality profile in Iran (29 provinces) over the years 2006 to 2010. Tehran: Ministry of Health and Medical Education; 2013.
5. Goodman MS, Li Y, Tiwari RC. Survival analysis with change point hazard functions. 2006.
6. Mazumdar M, Glassman JR. Categorizing a prognostic variable: review of methods, code for easy implementation and applications to decision-making about cancer treatments. Stat Med. 2000;19(5):513–32. [PubMed: 1062397].
7. Buettner P, Garbe C, Guggenmoos-Holzmann I. Problems in defining cutoff points of continuous prognostic factors: example of tumor thickness in primary cutaneous melanoma. J Clin Epidemiol. 1997;50(1):1201-10. [PubMed: 9393776].
8. Wang J, Liew G, Klein R, Rochtchina E, Khudtson MD, Klein BE, et al. Retinal vessel diameter and cardiovascular mortality: pooled data analysis from two older populations. Eur Heart J. 2007;28(16):9864–92. doi: 10.1093/eurheartj/ehm221. [PubMed: 17626032].
Tashima R, Nishimura R, Osako T, Nishiyama Y, Nakano M, et al. Evaluation of an Optimal Cut-Off Point for the Ki-67 Index as a Prognostic Factor in Primary Breast Cancer: A Retrospective Study. PLoS One. 2015;10(7). e0119565. doi: 10.1371/journal.pone.0119565. [PubMed: 26177501]. [PubMed Central: PMC4503758].

Mohsin SK, Weiss H, Havighurst T, Clark GM, Berardo M, Roanhole D, et al. Progesterone receptor by immunohistochemistry and clinical outcome in breast cancer: a validation study. Mod Pathol. 2004;17(12):1545–54. doi: 10.1038/modpathol.3800229. [PubMed: 15272777].

Lopez-Abente G, Ardanaz E, Torrella-Ramos A, Mateos A, Delgado-Sanz C, Chirlaque MD, et al. Changes in colorectal cancer incidence and mortality trends in Spain. Ann Oncol. 2010;21 Suppl 3:iii76–82. doi: 10.1093/annonc/mdq091. [PubMed: 20427364].

Othus M, Li Y, Tiwari R. Change point-cure models with application to estimating the change-point effect of age of diagnosis among prostate cancer patients. J Appl Stat. 2012;39(4):901-11. doi: 10.1080/02664763.2011.626849. [PubMed: 22544992]. [PubMed Central: PMC3337711].

Farewell VT. The use of mixture models for the analysis of survival data with long-term survivors. Biometrics. 1982;38(4):1041–6. [PubMed: 7168793].

Wang D. Survival analysis with long-term survivors. New York: Wiley; 2000.

Li CI, Malone KE, Porter PL, Daling JR. Epidemiologic and molecular risk factors for contralateral breast cancer among young women. Br J Cancer. 2003;89(3):513-8. doi: 10.1038/sj.bjc.6601042. [PubMed: 12888823]. [PubMed Central: PMC394484].

Lamont EB, Herndon J2, Weeks JC, Henderson IC, Earle CC, Schilsky RL, et al. Measuring disease-free survival and cancer relapse using Medicare claims from CALGB breast cancer trial participants (companion to 9344). J Natl Cancer Inst. 2006;98(18):1335-8. doi: 10.1093/jnci/djl305. [PubMed: 16985253]. [PubMed Central: PMC4158033].

Li Y, Tiwari RC, Guha S. Mixture cure survival models with dependent censoring. J Royal Statistic Soc Series B Statistic Methodol. 2007;69(3):285-306. doi: 10.1111/j.1467-9868.2007.00589.x.