RESEARCH ARTICLE

Editorial Process: Submission:11/11/2017 Acceptance:05/22/2018

Comparison Cure Rate Models by DIC Criteria in Breast Cancer Data

Ahmad Reza Baghestani[1], Parviz Shahmirzalou[2,3]*, Soheila Sayad[3], Mohammad Esmaeil Akbari[3], Farid Zayeri[4]

Abstract

Background: One of the malignant tumors is Breast Cancer (BC) that starts in the cells of breast. There is many models for survival analysis of patients such as Cox PH model, Parametric models etc. But some disease are that all of patients will not experience main event then usual survival model is inappropriate. In addition, In the presence of cured patients, if researcher can specify distribution of survival time, usually cure rate models are preferable to parametric models. Distribution of Survival time can be Weibull, Log normal, Logistic, Gamma and so. Comparison of Weibull, Log normal and Logistic distribution for finding the best distribution of survival time is purpose of this study. **Material and Methods:** Among 787 patients with BC by Cancer Research Center recognized and followed from 1985 until 2013. Variables stage of cancer, age at diagnosis, tumor size and Number of Removed Positive Lymph Nodes (NRPLN) for fitting Cure rate model were selected. The best model selected with DIC criteria. All analysis were performed using SAS 9.2. **Results:** Mean (SD) of age was 48.47 (11.49) years and Mean of survival time and Maximum follow up time was 326 and 55.12 months respectively. During following patients, 145 (18.4%) patients died from BC and others survived (censored). Also, 1-year, 5-year and 10-year survival rate was 94, 77 and 56 percent respectively. Log normal model with smaller DIC were selected and fitted. All of mentioned variables in the model were significant on cure rate. **Conclusion:** This study indicated that survival time of BC followed from Log normal distribution in the best way.

Keywords: Breast neoplasm- cancer research center- cure rate model- deviance information criteria (DIC)

Asian Pac J Cancer Prev, **19 (6)**, 1601-1606

Introduction

One of the malignant tumors is Breast Cancer (BC) that starts in the cells of breast. This disease is more commonly seen in women, but men can get it too (Wang et al., 2015). In 2012, there is approximately 14 million new cases of cancer in addition 8.2 million cancer related deaths. The five most common cancer among women were breast, colorectal, lung, cervix and stomach cancer. BC caused to 521,000 deaths in 2012 solely (Yi et al., 2005). In the same year, 1.7 million (12% of all new cancer cases) new cases of BC diagnosed (second most common cancer overall) that is 25% of all cancers in women (Tao et al., 2014).

Survival analysis is a collection of statistical procedures for data analysis for which the outcome variable of interest is time until an event occurs. Time can be days, weeks, months or years and event can be death, recurrence, incidence relapse from remission. In survival analysis, we usually refer to time variable as survival time (Kleinbaum and Klein, 2012).

There are two main regression models. One is Cox Proportional Hazards (PH) model as semi-parametric model and another is accelerated failure time model as parametric model. Baseline hazard in Cox PH model is not assumed to be of a parametric form and this model did not specify any distribution for time to event. Then this model is most used model for survival analysis (Pourhoseingholi et al., 2007). Efron (1977) and Oakes (1977) showed that under certain circumstances, parametric models are better than Cox PH model. In parametric model unlike cox model, survival time assumed to follow some distribution such as Normal, Poisson and Binomial distribution. Parametric models are better than cox model, when researcher can specify a suitable distribution to survival time.

Now, assume that the studied population formed from two groups. Patients who experience the event of interest (susceptible individuals) and cured patients that never experience it (non-susceptible individuals) (Maller and Zhou, 1996). With developments in medical sciences and new treatments, we now encounter with more disease that some patients are expected to be cured. Parametric models such as Weibull or Gamma do not account for

*[1]Physiotherapy Research Centre, Department of Biostatistics, Faculty of Paramedical Sciences, [3]Cancer Research Center, [4]Department of Biostatistics, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, [2]Social Determinants of Health Research Center, Yasuj University of Medical Sciences, Yasuj, Iran *For Correspondence: pshkhoei@gmail.com*

probability of cure (Cox and Oakes, 1984). Although subtle, there is difference between censoring and cure. A patient is censored if does not fail within the monitoring time window of a particular subject, while cure refers to one who will not fail within any reasonable monitoring time window (Gu et al., 2011).

There is two main models for population with cure and susceptible individuals. One is Mixture cure rate model or standard cure rate model. In this model supposed that there is π percent of individuals as cured and 1- π percent of individuals as susceptible. Survival function of this model written as:

$$S(t) = \pi + (1-\pi)S_0(t)$$

Where $\pi \in (0,1)$ is the mixing parameter (cure fraction) and $S_0(t)$ denotes survival function of susceptible patients (Mazucheli et al., 2013). Another cure rate model named as Non-mixture cure rate model. The survival function for Non-mixture cure rate model is:

$$S(t) = \pi^{1-S_0(t)}$$

Again, $\pi \in (0,1)$ and $S_0(t)$ is survival function for susceptible individuals (Martinez et al., 2013). One of ways for identifying existence of cure fraction is Kaplan-Meier curve. Then, Survival curve has a plateau among long-term survivors. Then cure rate model may be useful to analyze data (Baghestani et al., 2015). The aim of this study was to compare cure rate models to find the best model fitting survival time of BC data.

Materials and Methods

Data of 787 patients with BC were used which recognized by Cancer Research Center of Shahid Beheshti University, Tehran-Iran from 1985 until 2013. Then, Variables size of tumor, stage of cancer, age at diagnosis time and Number of Removed Positive Lymph Nodes (NRPLN) were imported to study that are clinically important too. All of variables are quantitative unless stage of cancer with 2 subgroups. In the present study, patients with 1 and 2 stage of cancer to group A and patients with stage 3 and 4 imported to group B.

Cure rate models to estimate the cure fraction was first developed by Boag in 1949 (Boag and John, 1949). There is two type of cure rate model, Mixture and Non-Mixture. In the present study, Non-Mixture cure rate model were fitted. This model formulated within a biological context presented by Yakovlev and Tsodikov (1996) then Chen et al., (1999). The progression (or promotion) time for the jth tumor cell is denoted by $R_j, j = 1,...,N$, where N is an unobservable (or latent) random variable indicating the number of clonogen tumor cells at the end of treatment which can produce a detectable cancer. In individuals with N=0, we can define $P(R_0 = \infty) = 1$ to represent a cure. In this study, latent variable has Negative binomial distribution with probability function such as:

$$p_m = P(N = m) = \frac{\Gamma(\eta^{-1}+m)}{\Gamma(\eta^{-1})m!}\left(\frac{\eta\theta}{1+\eta\theta}\right)^m (1+\eta\theta)^{-1/\eta},$$

Where m=0,1,2,...., $\eta > 0$, $\theta > 0$, $E(N) = \theta$, and $Var(N) = \theta + \eta\theta^2$ (Piegorsch, 1999; Saha and Paul, 2005; Cancho et al., 2011). The population survival time is indicated by:

$$S_{pop}(t) = \{1 + \eta\theta F(t)\}^{-1/\eta}$$

And the corresponding population hazard function is:

$$h_{pop}(t) = \{1 + \eta\theta F(t)\}^{-1/\eta - 1}\theta f(t)$$

Then, F(t) indicate Cumulative distribution function of survival time. Survival time is equal to the time until the patient's death or the censoring time. Covariates in the model connected to Θ by $E(N_i) = \theta_i = \exp(x_i\beta), i = 1,...,n$ where n is the number of subjects and β is equal to a k*1 vector of regression coefficient. It is apparent that there is a relation between covariates and cure rate. Since $p_0 = (1+\eta\theta)^{-1/\eta}$, we can conclude an inverse relation between them. Thus, in a covariate with positive coefficient, an increase in its value implies a reduction in the cure rate (keeping constant all other covariates) (Cancho et al., 2011).

Cure rate models with different distribution of survival time were fitted. Weibull, Log normal and Logistic distribution specified for survival time and specified models were fitted to data with SAS 9.2. Among interpreting results, models were compared together by Deviance Information Criterion (DIC). This criterion has been introduced by David Spiegelhalter in 2002:

$$DIC = -2\ln\{p(x \mid \theta)\} + p_D$$

Where P_D is the complexity as a measure of effective number of parameters, which is:

$$p_D = E[-2\ln\{p(x \mid \theta)\} + 2\ln\{p(x \mid \theta)\}]$$

P_D want to say that the effective number of parameters is the posterior mean deviance minus the deviance measured at the posterior mean of the parameters (Spiegelhalter et al., 2002).

All of cure rate models have latent variable with negative binomial distribution. Latent variable is defined as the number of remaining cancer cells in cancerous tissue after receiving treatment (Hanin et al., 2001; Cancho et al., 2015). Finally, results of fitting model with small DIC were reported.

Results

A total of 787 women with BC and mean (SD) of age equal to 48.47 (11.49) years were included in the analysis. The patients age at diagnosis BC ranged from 17 to 84 years with Mean of survival time and Maximum follow up time equal to 55.12 and 326 months respectively. During following patients, 145(18.4%) patients with mean of survival time equal to 46.07 months died from BC and others survived (censored). In average, every patient has a tumor of 3.23 cm and near 3 (exactly 2.80) number of removed positive lymph nodes. Also, stage of cancer has two subgroups which described in Table 1. Also, 1-year, 5-year and 10-year survival rate was 94, 77 and 56 percent respectively.

Drawing Kaplan-Meier curve is one of methods for finding existence of cure fraction and patients that cured. There are cured patients if curve be flat after following patients long time (Figure 1). It seems after 180 month,

Table 1. Baseline Characteristics of Patients by Stage of Cancer

Variable	Category	No (%)	Survival mean in month (SD)	Number of died(%)
stage of cancer	A	529 (67.2)	56.53 (45.96)	47 (8.9)
	B	258 (32.8)	52.25 (41.95)	98 (38)

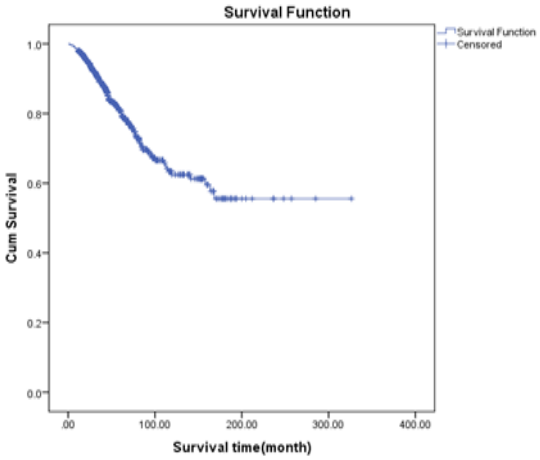

Figure 1. Kaplan-Meier Curve for Overall Cumulative Survival of Women with BC

curve has been flat and continued like a straight line forever.

Thus, Cur rate models with Weibull, Log normal and Logistics distributions specified for survival time were fitted and DIC criteria for all of them indicated in Table 2. This table showed that DIC of Log normal is smaller than others. Although, DIC of model with weibull distribution is near to that.

Table 2 indicated that DIC criteria of Log normal cure rate model is smaller than Weibull and Logistic model. Although results of three models were indicated in Table 3, most comparisons were done on Log normal cure rate model.

Table 2. DIC Criteria for Weibull, Log Normal and Logistic Cure Rate Model

Cure rate model	DIC
Weibull	1942.07
Log normal	1820.56
Logistic	2356.04

Result of fitting cure rate model with Log normal distribution for survival time showed that all of four variables are significant in cure rate of patients with BC. All of four variables, Size of tumor, NRPLN, age and stage 2 and 3 of cancer have mean (posterior summaries) with positive signs Then, cause to decrease probability of being cure in patients. For example, every 1 cm increase in size of tumor, probability of being cure decreased to 4%. Also, Cure Rate were calculated for every variable separately.

Although Weibull cure rate model has a little lager DIC than log normal, results of fitting this model indicated that Age at diagnosis, NRPLN and Size of tumor were significant on cure rate of patients. On the other side, DIC of Logistic cure rate model was more larger than two other

Table 3. Results of Fitting Cure Rate Models to Breast Cancer Data

Model	Parameters	Mean** (SD)	Credible Interval		Sig.	Cure rate
			%2.5	%97.5		
Weibull	Intercept	-2.07	-4.17	0.19	-	-
	stage of cancer A	REF	REF		-	-
	B	0.58	-1.70	2.11		0.97
	Age	0.05	0.006	0.10	*	0.99
	NRPLN	0.21	0.09	0.37	*	0.99
	tumor size(cm)	0.38	0.20	0.60	*	0.98
Log normal	Intercept	-1.40	-1.92	-0.79	-	-
	stage of cancer A	REF	REF		*	-
	B	1.00	0.38	1.46		0.85
	Age	0.04	0.03	0.06	*	0.99
	NRPLN	0.11	0.05	0.16	*	0.98
	tumor size(cm)	0.27	0.18	0.37	*	0.96
Logistic	Intercept	-7.04	-8.02	-5.53	-	-
	stage of cancer A	REF	REF		*	-
	B	2.16	1.66	2.88		0.99
	Age	0.07	0.04	0.09	*	0.99
	NRPLN	0.01	-0.05	0.6	-	0.99
	tumor size(cm)	0.18	0.08	0.27	*	0.99

*, significant; **, mean of posterior summaries; ***, shape or disperion parameter; ****, mean parameter

model, regardless that effect of Stage of cancer, Age and NRPLN were significant on cure rate.

Discussion

By last developments in medicine, some patients with enough follow up in time will not experience main event e.g. death, disease incidence, relapse from remission, recovery or any designated experience of interest. In fact, target population divided into two subgroup. One group unlike others are cured and is not susceptible to main event and will not experience main event forever (Scolas et al., 2016). When a fraction of population is not susceptible, the survival distribution is improper, leading the survival function to level off at a value different from zero. In this case, cure rate models is a suitable selection. There is two type of cure rate models. One is Mixture cure rate model (Boag, 1949; Berkson and Gage, 1952) and another Non-mixture cure rate model or Promotion time model (Tsodikov, 1998; Tsodikov et al., 2003). In the present study, Non-mixture Cure rate model with different distribution for survival time were fitted and the best model selected by DIC criteria. Since model with smaller DIC is better, cure rate model with Log normal distribution was selected. All of discussions were on results of Log normal cure rate model.

The relationship between stage of cancer and survival of patients with BC has been published in more studies (Gakwaya et al., 2008; Lee et al., 2007). Detecting BC in earlier stage cause to improve life expectancy and treatment of patients (Bray et al., 2004). In this study, Stage of cancer is a variable with 2 subgroup ((stage A (stage 1 and 2) and stage B (stage 3 and 4)). Stage A is referenced and stage B compared with them. Results showed that patients in stage B has less probability (15 percent) of being cure than patients in stage A. A Study by Carey showed that African-American cases had worse survival than Non African-American cases and variable stage of cancer was significant factor ($P<0.001$) in survival of patients with BC (Carey et al., 2006). Another study indicated that black women are diagnosed at a later stage of BC compared to white women. Also, this study conclude that patients with high educational level or income have low stage of cancer ($P<0.001$). Maybe, one reason is that patients with low income or education are less likely to participate in BC early detection programs (Wells and Horn, 1992). In importance of stage of cancer, Rahimazadeh et al studied 345 patients with BC. They fit cure rate model in a Non-mixed cure rate model with both Poisson and Negative Binomial distribution. They used Age at diagnosis time, metastasis and the stage of BC as prognosis factors in cure rate model. According to the results of fitting this model, metastasis and stage of BC were the significant factors in both model, but age at diagnosis was significant only in Negative Binomial model (Rahimzadeh et al., 2014).

Age at diagnosis time of BC is one of variables that role of them in cure fraction of BC patients was surveyed. This study indicated that age at diagnosis time is a factor determining probability of being cure. In fact, every one year increase in age at diagnosis time cause to 1% decrease in cure probability. In a study by Hill et al., (2015), indicated that high age at diagnosis time of BC will reduce survival of patients. Another study in Korea, estimated mean age at diagnosis equal to 48.3, very near this study (48.47 years old) (Lee and Oh, 2014). in a study by Brandt et al, age at diagnosis time categorized to 6 age group (<40, 40-49, 50-59, 60-69, 70-79, >80 years). Effect of age was surveyed in relation to survival, axillary lymph node involvement, diagnostic period and results of them reported. They conclude that young patients (<40 years) and old patients (>80 years) was positively associated with high BC specific mortality (BCSM). This conclusion is predominantly apparent for young women with negative auxiliary lymph node involvement BC. Finally, 80 years and more age at diagnosis time, independent of stage and diagnosis period, is a risk factor for BCSM (Brandt et al., 2015).

More study, surveyed effect of lymph node involvement in survival of BC patients (Solak et al., 2015; Giuliano et al., 1994; Albertini et al., 1996; Carter et al., 1989). In the present study, NRPLN detected as a significant factor in cure rate of patients. Every one more NRPLN cause to 2% decrease in probability of being cure. In average, every patient has 3 NRPLN. In a study, Krag et al categorized number of nodes positive (0, 1-3), number of nodes removed (<10, ≥10) and age group (40-49, 50-79 years old). They indicated that even when all regional lymph nodes are pathologically negative, the number of nodes removed is significantly associated with survival. Also, concluded that in the group with 1-3 pathologically positive nodes, the number of nodes was associated with an even greater survival difference than with the node-negative group (Kreg and Single, 2003).

Size of tumor is one of the important factors for BC (Tabar et al., 1992; Tabar et al., 2000; Tubiana and Koscielny, 1991; Koscielny et al., 1984; Tubiana and Koscielny, 1999). Mean size of tumor in present study population was 3.23 cm. The effect of this variable on cure rate was significant. Study indicated that every 1 cm increase in size of tumor will decrease probability of being cured equal 4%. Rosenberg et al surveyed the effect of tumor size, grade, race and year at diagnosis time on survival of BC patients by Cox PH regression model. They conclude that all of variables have effect on disease-specific survival BC but effect of age at diagnosis time and stage of cancer vary over time. Finally, they indicated that large tumor size and higher tumor grade have a large negative effects on survival (Rosenberg et al., 2005). Michaelson et al in a study tried to predicting survival of patients with BC using tumor size. This study as shown in many studies revealed that survival of patients decreases as size of tumor become larger. In other words, survival of BC patients related with tumor size directly and independent of the method of detection (Michaelson et al., 2002).

In the present study, 1 , 5 and 10-years survival rate of patients was found 94, 77 and 56 percent respectively. Fouladi and coworkers studied survival rate of patients with BC in Ardabil province of Iran. In this tudy, 5-year survival rate was found 51% (Fouladi et al., 2012). Another study by Rezaianzadeh et al., (2009)

that conducted in Southern of Iran, three and five-year overall survival rate were found to be 76 and 58 percent respectively. In the present study, five-year survival rate were found more than mentioned studies.

In Figure 1, we can find that after 180 month (15 years) Kaplan-Meier curve of BC patients be flat. This means that after mentioned time, patients be cured and risk of death from BC in patients be equal to general population. In a study, Haghighat et al. concluded that the observed mortality rate of BC patients are equal to expected mortality rate of general population after 15 years from diagnosis BC (Haghighat et al., 2013). There is a few limitation running this model. Among them, we mention categorizing stage of cancer to two stage, selecting three popular survival distribution from several of them and inaccessibility to dossier of some patients that caused to reduce sample size of study.

In conclusion, cure rate model can be used in presence of cured patients for surveying survival of patients and factors affecting cure rate specially BC disease. In cure rate models like parametric models, a suitable distribution specified to survival time of data. It is apparent that this specification affected results of models. Some of distributions for survival time of BC are Weibull, Log-normal, Logistic and Gamma. In presence of cured patient and ability to detect and specify accurate distribution to survival time, cure rate models are preferable to usual survival models. Cure rate model with Log normal distribution for survival time (and Negative Binomial distribution for latent variable) has better results clinically than Weibull and Logistic distribution.

Acknowledgements

Hereby I extend my appreciation to my dear colleagues of Biostatistics Group and Cancer Research Center of Shahid Beheshti University of Medical Sciences. This paper is a part of a thesis in Biostatistics Department of paramedical sciences faculty of Shahid Beheshti University of Medical Sciences.

References

Albertini JJ, Lyman GH, Cox C, et al (1996). LYmphatic mapping and sentinel node biopsy in the patient with breast cancer. *JAMA*, **276**, 1818-22.

Baghestani AR, Zayeri F, Akbari ME, et al (2015). Fitting cure rate model to breast cancer data of cancer research center. *Asian Pac J Cancer Prev*, **16**, 7923-7.

Berkson J, Gage RP (1952). Survival curve for cancer patients following treatment. *J Am Stat Assoc*, **47**, 501-15.

Boag JW (1949). Maximum likelihood estimates of the proportion of patients cured by cancer therapy. *J R Stat Soc Series B*, **11**, 15-53.

Boag, John W (1949). Maximum likelihood estimates of the proportion of patients cured by cancer therapy. *J R Stat Soc Series B Stat Methodol*, **11**, 15-53.

Brandt J, Garne JP, Tengrup I, Manjer J (2015). Age at diagnosis in relation to survival following breast cancer: a cohort study. *World J Surg Oncol*, **13**, 1-11.

Bray F, McCarron P, Parkin DM (2004). The changing global patterns of female breast cancer incidence and mortality. *Breast Cancer Res*, **6**, 229-39.

Cancho VG, Rodrigues J, de Castro Mr (2011). A flexible model for survival data with a cure rate: a Bayesian approach. *J Appl Stat*, **38**, 57-70.

Cancho VG, Rodrigues J, de Castro Mr (2015). A flexible model for survival data with a cure rate: a Bayesian approach. *J Appl Stat*, **38**, 57-70.

Carey LA, Perou CM, Livasy CA, et al (2006). Race, breast cancer subtypes, and survival in the carolina breast cancer study. *JAMA*, **295**, 2492-502.

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

Chen MH, Ibrahim JG, Sinha DA (1999). new Bayesian model for survival data with a surviving fraction. *J Am Stat Assoc*, **94**, 909–19.

Cox DR, Oakes D (1984). Analysis of Survival Data, The scope of survival analysis, Chapman and Hall, London, pp 1-11.

Efron B (1977). The effciency of Cox's likelihood function for censored data. *J Am Stat Assoc*, **72**, 557–65.

Fouladi N, Amani F, Harghi AS, Nayebyazdi N (2011). Five year survival of women with breast cancer in Ardabil, north-west of Iran. *Asian Pac J Cancer Prev*, **12**, 1799-801.

Gakwaya A, Kigula-Mugambe JB, Kavuma A, et al (2008). Cancer of the breast: 5-year survival in a tertiary hospital in Uganda. *Br J Cancer*, **99**, 63-7.

Giuliano AE, Kirgan DM, Guenther JM, Morton DL (1994). Lymphatic mapping and sentinel lymphadenectomy for breast cancer. *Ann Surg*, **220**, 391-401.

Gu Y, Sinha D, Banerjee S (2011). Analysis of cure rate survival data under proportional odds model. *Lifetime Data Anal*, **17**, 123-34.

Haghighat S, Akbari ME, Ghaffari S, Yavari P (2012). Standardized breast cancer mortality rate compared to the general female population of Iran. *Asian Pac J Cancer Prev*, **13**, 5525-8.

Hanin LG, Zaider M, Yakovlev AY (2001). Distribution of the number of clonogens surviving fractionated radiotherapy: a long-standing problem revisited. *Int J Radiat Biol*, **77**, 205-13.

Hill J, Watanabe-Galloway S, Shostrom V, Nsiah-Kumi P (2015). Breast cancer survival among African-Americans living in the Midwest: Disparities and recommendations to decrease mortality. *J Natl Black Nurses Assoc*, **26**, 8-14.

Kleinbaum DG, Klein M (2012). Survival analysis: a self-learning text. Introduction to survival analysis. Springer, New York, 3rd ed, pp 4.

Koscielny S, Tubiana M, Lê MG, et al (1984). Breast cancer: relationship between the size of the primary tumor and the probability of metastatic dissemination. *Br J Cancer*, **49**, 709 –15.

Krag DN, Single RM (2003). Breast cancer survival according to number of nodes removed. *Ann Surg Oncol*, **10**, 1152-9.

Lee J, Oh M (2014). The Korean breast cancer S. effects of interval between age at first pregnancy and age at diagnosis on breast cancer survival according to menopausal status: a register-based study in Korea. *BMC Womens Health*, **14**, 113.

Lee J-H, Yim S-H, Won Y-J, et al (2007). Population-based Breast Cancer Statistics in Korea during 1993-2002: Incidence, Mortality, and Survival. *J Korean Med Sci*, **22**, 11-6.

Maller RA , Zhou X (1996). Survival analysis with long-term survivors, formulating tests for the presence of immunes and sufficient follow-up. John Wiley and Sons, pages 1-10.

Martinez EZ, Achcar JA, J‡come AAA, Santos JS (2013). Mixture and non-mixture cure fraction models based on the generalized modified Weibull distribution with an application to gastric cancer data. *Lifetime Data Anal*, **112**,

343-55.
Mazucheli J, Coelho-Barros EA, Achcar JA (2013). The exponentiated exponential mixture and non-mixture cure rate model in the presence of covariates. *Comput Methods Programs Biomed*, **112**, 114-24.
Michaelson JS, Silverstein M, Wyatt J, et al (2002). Predicting the survival of patients with breast carcinoma using tumor size. *Cancer*, **95**, 713-23.
Oakes D (1977). The asymptotic information in censored survival data. *Biometrika*, **64**, 441–8.
Piegorsch WW (1990). Maximum likelihood estimation for the negative binomial dispersion parameter. *Biometrics*, **46**, 863–7.
Pourhoseingholi MA, Hajizadeh E, Moghimi Dehkordi B, et al (2007). Comparing Cox regression and parametric models for survival of patients with gastric carcinoma. *Asian Pac J Cancer Prev*, **8**, 412-6
Rahimzadeh M, Baghestani AR, Gohari MR, Pourhoseingholi MA (2014). Estimation of the cure rate in Iranian breast cancer patients. *Asian Pac J Cancer Prev*, **15**, 4839-42.
Rezaianzadeh A, Peacock J, Reidpath D, et al (2009). Survival analysis of 1148 women diagnosed with breast cancer in Southern Iran. *BMC Cancer*, **9**, 168.
Rosenberg J, Chia YL, Plevritis S (2005). The effect of age, race, tumor size, tumor grade, and disease stage on invasive ductal breast cancer survival in the U.S. SEER database. *Breast Cancer Res Treat*, **89**, 47-54.
Saha K, Paul S (2005). Bias-corrected maximum likelihood estimator of the negative binomial dispersion parameter. *Biometrics*, **61**, 179–85.
Scolas S, El Ghouch A, Legrand C, Oulhaj A (2016). Variable selection in a flexible parametric mixture cure model with interval-censored data. *Stat Med*, **35**, 1210-25.
Solak M, Turkoz FP, Keskin O, et al (2015). The lymph node ratio as an independent prognostic factor for non-metastatic node-positive breast cancer recurrence and mortality. *J Buon*, **20**, 737-45.
Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A (2002). Bayesian measures of model complexity and fit. *J R Stat Soc Series B Stat Methodol*, **64**, 583-639.
Tabar L, Dean PB, Duffy SW, Chen HH (2000). A new era in the diagnosis of breast cancer. *Surg Oncol Clin North Am*, **9**, 233–77.
Tabar L, Fagerberg G, Duffy SW, et al (1992).Update of the Swedish two-county program of mammo-graphic screening for breast cancer. *Radiol Clin North Am*, **9**, 233–77.
Tao Z, Shi A, Lu C, et al(2015). Breast cancer: epidemiology and etiology. *Cell Biochem Biophys*, **72**, 333-8.
Tsodikov A (1998). A proportional hazards model taking account of long-term survivors. *Biometrics*, **54**, 1508-16.
Tsodikov AD, Ibrahim JG, Yakovlev AY (2003). Estimating cure rates from survival data: An alternative to two-component mixture models. *J Am Stat Assoc*, **98**, 1063-78.
Tubiana M, Koscielny S (1991). Natural history of human breast cancer: recent data and clinical implications. *Breast Cancer Res Treat*, **18**, 125–40.
Tubiana M, Koscielny S (1999). The rationale for early diagnosis of cancer—the example of breast cancer. *Acta Oncol*, **38**, 295–303.
Wang LW, Yang GF, Chen JM, et al (2014). A clinical database of breast cancer patients reveals distinctive clinico-pathological characteristics: a study from central China. *Asian Pac J Cancer Prev*, **15**, 1621-6.
Wells BL, Horm JW (1992). Stage at diagnosis in breast cancer: race and socioeconomic factors. *Am J Public Health*, **82**, 1383–5.
Yakovlev AY, Tsodikov AD (1996). Stochastic models of tumor latency and their biostatistical appli-cations. World Scientific, New Jersey. **18**, pp 3156-7.
Yi T, Kabha E, Papadopoulos E, Wagner G (2014). 4EGI-1 targets breast cancer stem cells by selective inhibition of translation that persists in CSC maintenance, proliferation and metastasis. *Oncotarget*, **5**, 6028-37.