Assessment of the fitness of Cox and parametric regression models of survival distribution for Iranian breast cancer patients' data

Article · January 2019

DOI: 10.4103/japtr.JAPTR_360_18

11 authors, including:

Reza Shekarriz
Shahid Beheshti University of Medical Sciences

Parastoo Amiri
Shahid Beheshti University

Marjan Vejdani
Sabzevar University of Medical Sciences

Pezhman Farshidmehr
Tehran University of Medical Sciences

Original Article

Assessment of the fitness of Cox and parametric regression models of survival distribution for Iranian breast cancer patients' data

Maryam Mohseny, Reza Shekarriz-Foumani[1], Parastoo Amiri[2], Marjan Vejdani[2], Pezhman Farshidmehr[3], Hossein Zabihi Mahmoudabadi[4], Farzaneh Amanpour[5], Pegah Mohaghegh[6], Farzad Tajdini, Azadeh Sayarifard[7], Esmat Davoudi-Monfared[8]

Departments of Health and Community Medicine and [1]Department of Community Medicine, Medical School, Shahid Beheshti University of Medical Sciences,Tehran, [2]Iranian Research Center on Healthy Aging, Sabzevar University of Medical Sciences, Sabzevar, [3]Department of Vascular Surgery, Sina Hospital, Tehran University of Medical Sciences, [4]Department of Surgery, School of Medicine, Sina Hospital, Tehran University of Medical Sciences,Tehran, [5]Department of Epidemiology and Biostatistics, School of Public Health, Shahroud University of Medical Sciences, Shahroud, [6]Department of Community Medicine, School of Medicine, Arak University of Medical Sciences, Arak, Iran, [7]Community Based Participatory Research Center, Iranian Institute for Reduction of High-Risk Behaviors, Tehran University of Medical Sciences, [8]Health Management Research Center & Department of Community Medicine, Faculty of Medicine, Baqiyatallah University of Medical Sciences, Tehran, Iran

J. Adv. Pharm. Technol. Res.

Abstract

Factors affecting the time of survival after breast cancer (BC) diagnosis remain unknown. However, some of the prognostic factors have been identified. The aim of this study was to investigate the effects of biologic and socioeconomic factors on long-term survival of BC patients. This was a descriptive chart review and survey of all women with a confirmed diagnosis of BC registered in Shohada-e-Tajrish Cancer Research Center database from March 2004 to March 2015. The checklist of study consisted of biologic, demographic, reproductive, genetic, medical, and therapeutic information of patients. The minimum time of follow-up was 3 years and the maximum was 10 years. We then evaluated possible associations of these variables with BC survival using Cox and parametric regression models of survival analysis. The study population was 1276 BC patients. Their mean survival was 23 (range 1–120) months. Between the parametric models, Weibull regression model demonstrated the lowest Akaike information criterion and thus the best fit, and tumor size, number of lymph nodes, BC stage, educational level, and high-fat diet were significant in this model. Based on our findings, educational level, consumption of fat, and characteristics of tumor at the time of diagnosis (disease stage, tumor size, number of involved lymph nodes) are the most important prognostic factors affecting long-term survival of BC patients. We suggest that future studies assess the efficacy of possible interventions for these factors.

Key words: Breast cancer, hazard models, Iran, survival

Address for correspondence:
Dr. Esmat Davoudi-Monfared,
Department of Community Medicine, Health Management Research Center, Faculty of Medicine, Baqiyatallah University of Medical Sciences, Tehran, Iran.
E-mail: davoudimonfared@gmail.com

Access this article online	
Quick Response Code:	Website: www.japtr.org
	DOI: 10.4103/japtr.JAPTR_360_18

INTRODUCTION

Breast cancer (BC) is the most common cancer in developed countries, roughly affecting one in every 10 females. It

How to cite this article: Mohseny M, Shekarriz-Foumani R, Amiri P, Vejdani M, Farshidmehr P, Mahmoudabadi HZ, *et al*. Assessment of the fitness of Cox and parametric regression models of survival distribution for Iranian breast cancer patients' data. J Adv Pharm Technol Res 2019;10:39-44.

constitutes nearly 28% of cancers in women[1] and ranks second only after lung cancer as the most frequent cause of cancer deaths in women.[2]

Over the past few decades, the survival rate of BC patients has improved after diagnosis. Between 1999 and 2005, the 5-year survival rate among women who were diagnosed with cancer was 90%, compared with 1975–1977, when this being 75%.[3] Moreover, in 2017, the average 5-year survival rate for people with BC was 90%. The average 10-year survival rate is 83%.[4]

According to the current evidence, long survival time has improved due to advances in BC treatment approach and implementation of population-based screening programs of women at risk of BC. Unfortunately, all of women population does not show the same increase in survival. This difference in increasing the survival rate of women may depend on different factors. Many studies have examined these factors.

The etiology of BC remains unknown, but a broad range of factors have been identified that contribute to its high prevalence. These include biologic factors (age, family history, and ethnicity), social determinants of health (childhood conditions, social status, and educational level),[5] lifestyle-related factors (alcohol consumption, obesity, lack of physical exercise, stress, and smoking), and disease-related factors (stage at diagnosis, tumor grade, and comorbid conditions).[6] In addition, higher incidence has been commonly reported in women with higher socioeconomic status (SES), which may contribute to differences in lifestyle, family planning, and access to healthcare services.[7] On the other hand, women with higher SES survive more than the lower group, paradoxically.[8] The most important factors of long survival time in women with BC were socioeconomic factors such as regional variations, race/ethnicity, SES, and urbanization in previous studies.[9-11]

Nevertheless, evidence is inconclusive, and controversial results still exist.[12,13] Specifically, studies evaluating associations of individual factors as well as community determinants of SES with BC survival are lacking. We aimed to investigate the effects of biological and lifestyle factors on BC survival using parametric regression models of survival distribution analysis.

MATERIALS AND METHODS

In this descriptive chart review, we surveyed all women who had a record in the Shohada-e-Tajrish Cancer Research Center database with a confirmed diagnosis of BC, from March 2004 to March 2015. Our inclusion criteria were a confirmed BC diagnosis, living in or in proximity of Tehran, and having initiated treatment after the diagnosis. Women with missing data were excluded. The minimum number of patients required for this study at an alpha error of 5% and a beta error of 20% was 380.

We designed a specific checklist for this study. We considered years of education as a criterion for SES since it has been shown that it is a valid and reliable determinant of SES in social health studies in Iran. Our biologic and demographic variables included age, time of diagnosis, education level, marital status, fat content of diet, and smoking. Reproductive variables included gravidity, parity, miscarriages/abortions, breastfeeding duration, hormone consumption, duration of hormone consumption, age at first menstrual period, and age at menopause. Disease-related factors included disease stage, tumor grade, final pathology, and lymph node involvement. Genetic variables included family history of BC, estrogen receptor (ER) status, progesterone receptor (PR) status, and HER2, P53, and KI67 gene mutations. Finally, treatment-related factors included receiving chemoradiotherapy, hormonal therapy, and herceptin, as well as recurrence/metastasis.

Survival was evaluated during the aforementioned 10-year period with 3 years being the minimum acceptable duration of follow-up. Age, disease stage, and similar variables were assessed using Kaplan–Meier curves. Significant variables were then entered into the Cox and parametric regression models to investigate their correlation with survival outcomes.

RESULTS

This study population consisted of 1276 BC patients. The mean survival time was 23 (range 1–120) months. Table 1 presents the demographic characteristics of the study participants.

Tumor size, number of involved lymph nodes, BC stage, tumor grade, PR positivity, lymphovascular invasion, educational level, and high-fat diet were statistically significant in Kaplan–Meier modeling and were selected for Cox and parametric regression models. We excluded tumor grade in this step because of its overlap with tumor size and included age in all of our models because of its previously established importance although it failed to show significance in the Kaplan–Meier model in our study.

Assumptions of Cox model were analyzed using Schoenfeld residuals method, which approved the assumption of similarity of hazard ratios (HRs) for all variables. This showed that the risk of death was similar between different treatment groups in time, enabling the calculation of a constant for its effect.

Table 1: Case summary of patients with breast cancer

Variables	Total (*n*)	Event (%)
Age (years)		
<50	1048	79 (7.5)
>50	765	55 (7.2)
Stage		
1	349	9 (2.6)
2	731	34 (4.7)
3	466	49 (10.5)
4	61	23 (37.7)
ER		
No	826	32 (3.9)
Yes	601	55 (9.2)
PR		
No	486	39 (8)
Yes	1027	55 (5.4)
Education		
Illiterate	87	18 (20.7)
Elementary	377	42 (11.1)
Diploma	532	36 (4.9)
Academic	466	22 (4.7)
Rich food		
No	594	19 (3.2)
Yes	684	53 (7.7)
Side		
Left	313	11 (4.5)
Right	315	11 (3.5)
Abortion		
No	1031	76 (7.3)
Yes	573	32 (5.6)
Breastfeeding (months)		
0	258	15 (5.8)
1-24	324	20 (6.2)
>24	948	63 (6.6)
Total number of patients	1960	137 (7)
Side		
Left	313	11 (4.5)
Right	315	11 (3.5)
Right and left	10	0
Grade		
1	193	7 (3.6)
2	813	46 (5.7)
3	489	48 (9.8)
ER		
No	421	36 (8.6)
Yes	1100	59 (5.4)
HER20		
No	1011	53 (5.2)
Yes	151	5 (3.3)
P53		
No	266	15 (5.6)
Yes	156	20 (12.8)
Pregnancy		
0	209	13 (6.2)
1-4	1120	63 (5.6)
>4	311	32 (10.3)
Marital		
Single	111	4 (3.6)
Married	1505	105 (7)
Widow	40	0
Family history		
No	1314	85 (6.5)
Yes	218	12 (5.5)
Hormone		
No	964	59 (6.1)
Yes	405	24 (5.9)
Smoking		
No	1312	66 (5)
Yes	63	5 (7.9)

ER: Estrogen receptor, PR: Progesterone receptor

Risk distribution of the mortality of BC was determined using Weibull regression, lognormal, log-logistic, and Gompertz models. The goodness of fit of these models was then tested by fragility analysis, which evaluates the changes of survival over time not accounted for by the variables in the model. All models were compared by Akaike information criterion (AIC), where lower AIC shows better fit.

All parametric models showed a lower AIC than the Cox model. Moreover, between the parametric models, Weibull regression model demonstrated the lowest AIC and thus the best fit. Different models are compared in Table 2.

We further investigated the goodness of fit using Cox-Snell residuals, which also confirmed the superiority of the Weibull model [Figure 1]. Risk distribution curve is illustrated in Figure 2.

Fragility model did not show any statistical significance in any of the parametric regression models.

The effect estimates were presented by HRs in proportional hazard models and by survival time ratios in accelerated failure time models. Tumor size, number of lymph nodes, BC stage, educational level, and high-fat diet were significant in the Weibull regression model.

Patients with a larger tumor (HR = 1.20), stage 4 BC (vs. stage 1) (HR = 9.84), and high-fat diet (HR = 2.7) had higher risk of mortality, while patients with a high-school diploma or higher level of education had lower mortality risk compared with those with lower educational levels.

Table 2: Fitness of survival models in patients with breast cancer

Variables	HR		TR			HR
	Cox (PH)	Weibull (PH)	Weibull (AFT)	Lognormal (AFT)	Log-logistic (AFT)	Gompertz (PH)
Tumor size	1.20* (1.04, 1.39)	1.20* (1.04, 1.39)	0.87* (0.79, 0.98)	0.85* (0.74, 0.99)	0.87* (0.78, 0.98)	1.19* (1.03, 1.39)
Number of lymph node	1.06** (1.02, 1.09)	1.06** (1.02, 1.09)	0.96** (0.94, 0.99)	0.95* (0.91, 0.99)	0.96** (0.93, 0.99)	1.05** (1.01, 1.09)
Stage						
1 (ref)	-	-	-	-	-	-
2	0.78 (0.15, 4.06)	0.79 (0.15, 4.14)	1.17 (0.37, 3.71)	1.43 (0.46, 4.43)	1.25 (0.41, 3.83)	0.81 (0.16, 4.23)
3	1.50 (0.27, 8.25)	1.46 (0.27, 7.99)	0.76 (0.23, 2.51)	0.93 (0.28, 3.17)	0.76 (0.24, 2.38)	1.52 (0.28, 8.29)
4	10.42* (1.66, 65.45)	9.84* (1.57, 61)	0.20* (0.05, 0.74)	0.22 (0.05, 1.05)	0.21* (0.06, 0.78)	10.22* (1.63, 63.8)
ER						
No (ref)	-	-	-	-	-	--
Yes	1.61 (0.72, 3.61)	1.58 (0.70, 3.53)	0.73 (0.41, 1.28)	0.82 (0.41, 1.66)	0.75 (0.42, 1.33)	1.60 (0.72, 3.6)
PR						
No (ref)	-	-	-	-	-	-
Yes	0.63 (0.28, 1.41)	0.62 (0.27, 1.39)	1.369 (0.79, 2.43)	1.05 (0.51, 2.18)	1.25 (0.71, 2.21)	0.61 (0.27, 1.36)
Education						
Illiterate (ref)	-	-	-	-	-	-
Elementary	0.34 (0.11, 1.04)	0.36 (0.12, 1.08)	2.05 (0.91, 4.62)	2.53 (0.82, 7.80)	2.29 (0.98, 5.34)	0.38 (0.12, 1.16)
Diploma	0.09** (0.02, 0.39)	0.10** (0.02, 0.39)	4.96** (1.73, 14.16)	6.22** (1.75, 22.08)	5.10** (1.81, 14.35)	0.10** (0.03, 0.412)
Academic	0.15** (0.04, 0.59)	0.16** (0.04, 0.62)	3.54* (1.30, 9.64)	4.77* (1.35, 16.83)	3.76** (1.41, 10.02)	0.17* (0.04, 00.65)
Rich food						
No (ref)	-	-	-	-	-	-
Yes	2.73* (1.06, 7.03)	2.7* (1.06, 6.91)	0.5* (0.25, 0.98)	0.6 (0.28, 1.28)	0.53 (0.27, 1.04)	2.74* (1.07, 7)
Age						
<50 (ref)	-	-	-	-	-	-
≥50	0.66 (0.26, 1.66)	0.68 (0.27, 1.70)	1.30 (0.69, 2.47)	0.86 (0.41, 1.81)	1.07 (0.58, 1.98)	0.67 (0.26, 1.67)
AIC	304.08	235.08	235.08	249.1	236.01	235.37

*Significant at α=0.05, **Significant at α=0.01. HR: Hazard ratio, TR: Survival time ratio, Ref: Baseline category. ER: Estrogen receptor, PR: Progesterone receptor, PH: Proportional hazard, AFT: Accelerated failure time, AIC: Akaike information criterion

DISCUSSION

Between the parametric models of survival analysis, Weibull regression model demonstrated the lowest AIC and thus the best fit.

In the present study, we evaluated the relationship between long-time survival of BC and various biologic and demographic (age, time of diagnosis, education level, marital status, fat content of diet, and smoking), reproductive (gravidity, parity, miscarriages/abortions, breastfeeding duration, hormone consumption, duration of hormone consumption, age at first menstrual period, and age at menopause), genetic (included family history of BC, ER status, PR status, and HER2, P53, and KI67 gene mutations), disease (disease stage, tumor grade, final pathology, and lymph node involvement), and treatment-related (receiving chemoradiotherapy, hormonal therapy, and herceptin, recurrence/metastasis) factors.

We found that educational level, high-fat diet, and disease characteristics such as cancer stage, tumor size, and number of positive lymph nodes were important prognostic factors that influenced survival.

The effect of educational level on BC survival is a controversial topic. While previous studies demonstrated higher incidence of BC in more educated patients,[14]

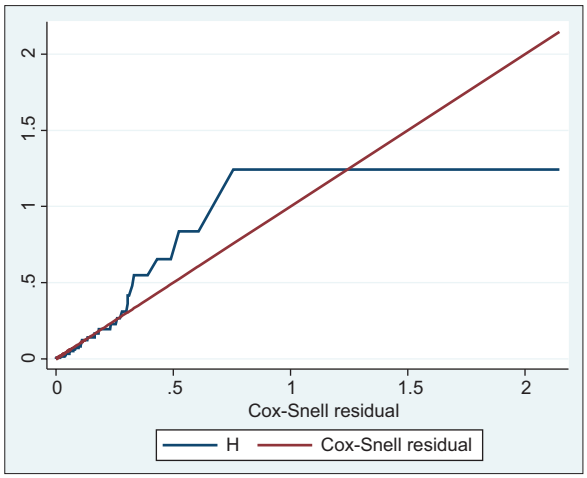

Figure 1: Cox-Snell residue for Weibull distribution in patients with breast cancer

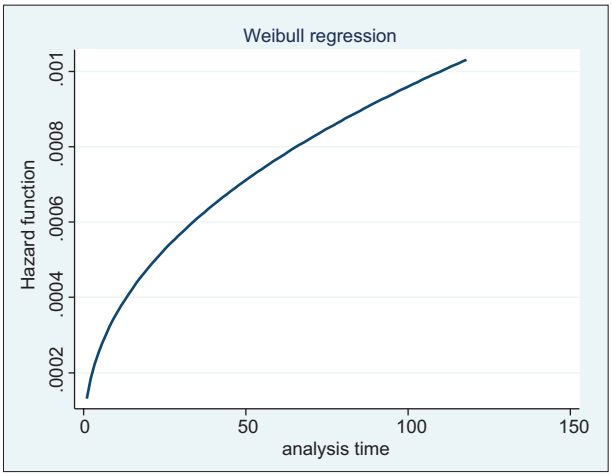

Figure 2: The risk distribution curve of the Weibull in patients with breast cancer

better survival rates have been reported in this patient population,[15] similar to our findings. In Herndon *et al.*'s study, the level of education below high-school diploma was a risk factor for death in women with BC.[16]

The level of education is closely related to the SES of individuals. Therefore, the difference in the level of education shows the different social and economic status of individuals. Higher BC incidence rates in educated patients may be explained by other SES components such as reproductive behavior, number of children, older age at first pregnancy, and shorter duration of breastfeeding.[17] Better survival, on the other hand, can be due to the fact that these patients are more inclined toward BC screening programs.[18,19] This inclination, in turn, would also translate into higher incidence rates in these patients, but at earlier stages, enabling more effective treatment of BC and hence better survival.

Another possible explanation of this difference in survival could be due to lead-time bias in studies in which educated patients are found to have better survival, while in fact the earlier diagnosis of their condition results in a false notion of better survival because of the longer period until death. In addition, women with lower SES tend to resist the complementary axillary surgery and adjuvant chemotherapy, which creates a poor outcome in treatment and can affect survival.[20] Better coverage by health insurance and thus accessibility to healthcare services are another possible advantages in educated patients, whereas those with lower educational levels may geographically and economically struggle to have such access to receive the most effective treatments.[21]

There is strong evidence that obesity is a potent risk factor for both cancer development and the prognosis of BC which is defined by anthropometric measurements and body mass index (BMI).[22]

We demonstrated that high-fat diet is another prognostic factor in BC patients.

This is similar to previous findings showing a direct relationship between fat consumption and recurrence and survival rates in BC patients. The studies suggested lower intake of saturated fat in the diet of BC patients in postdiagnosis period was associated with improved survival after BC diagnosis.[23,24] A meta-analysis by Xing *et al.*, moreover, showed that following a low-fat diet in BC patients reduced recurrence rate and in turn increased survival.[25] This relationship has been observed in both ER+ and ER− tumors, which may be explained by attenuated inflammatory processes and lower levels of inflammatory mediators in tissues. On the other hand, many studies have evaluated the effect of diet on incidence of BC, but the evidence is still inconclusive.[26] While the effect of Western high-fat diet is widely studied, we showed that an equivalent high-fat Eastern diet also adversely affected the survival in BC patients. This finding supports the notion that the fat content is the major player irrespective of type of diet. Having ruled out reproductive behavior in our study as a possible influence, we may conclude that diet could be considered the most important lifestyle element for survival in BC patients.

Besides education and diet, we emphasized the significant effects of disease characteristics such as stage, tumor size, and number of involved lymph nodes on survival. Many studies have shown the effect of tumor size independently of the number of lymph nodes on mortality rates in BC; our study further confirmed this finding.[27]

CONCLUSION

Among the various biologic, demographic, genetic, reproductive, and pathophysiologic factors associated with BC, the most important factors affecting long-term

survival included educational level, high-fat diet, and disease characteristics at the time of diagnosis. We suggest that future studies focus on interventions on these factors and evaluate their effectiveness regarding patient survival.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

REFERENCES

1. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JW, Comber H, *et al.* Cancer incidence and mortality patterns in Europe: Estimates for 40 countries in 2012. Eur J Cancer 2013;49:1374-403.
2. DeSantis CE, Fedewa SA, Goding Sauer A, Kramer JL, Smith RA, Jemal A, *et al.* Breast cancer statistics, 2015: Convergence of incidence rates between black and white women. CA Cancer J Clin 2016;66:31-42.
3. Horner MJ, Ries LA, Krapcho M, Neyman N, Aminou R, Howlader N, *et al.* SEER Cancer Statistics Review, 1975-2006. Bethesda, MD: National Cancer Institute; 1975-2006.
4. DeSantis CE, Ma J, Goding Sauer A, Newman LA, Jemal A. Breast cancer statistics, 2017, racial disparity in mortality by state. CA Cancer J Clin 2017;67:439-48.
5. Davoudi Monfared E, Mohseny M, Amanpour F, Mosavi Jarrahi A, Moradi Joo M, Heidarnia MA, *et al.* Relationship of social determinants of health with the three-year survival rate of breast cancer Asian Pac J Cancer Prev 2017;18:1121-6.
6. Davies NJ, Batehup L, Thomas R. The role of diet and physical activity in breast, colorectal, and prostate cancer survivorship: A review of the literature. Br J Cancer 2011;105 Suppl 1:S52-73.
7. Akinyemiju TF, Genkinger JM, Farhat M, Wilson A, Gary-Webb TL, Tehranifar P, *et al.* Residential environment and breast cancer incidence and mortality: A systematic review and meta-analysis. BMC Cancer 2015;15:191.
8. Warner ET, Tamimi RM, Hughes ME, Ottesen RA, Wong YN, Edge SB, *et al.* Racial and ethnic differences in breast cancer survival: Mediating effect of tumor characteristics and sociodemographic and treatment factors. J Clin Oncol 2015;33:2254-61.
9. Parise CA, Caggiano V. Regional variation in disparities in breast cancer specific mortality due to race/ethnicity, socioeconomic status, and urbanization. J Racial Ethn Health Disparities 2017;4:706-17.
10. Robert SA, Strombom I, Trentham-Dietz A, Hampton JM, McElroy JA, Newcomb PA, *et al.* Socioeconomic risk factors for breast cancer: Distinguishing individual- and community-level effects. Epidemiology 2004;15:442-50.
11. Sineshaw HM, Gaudet M, Ward EM, Flanders WD, Desantis C, Lin CC, *et al.* Association of race/ethnicity, socioeconomic status, and breast cancer subtypes in the national cancer data base (2010-2011). Breast Cancer Res Treat 2014;145:753-63.
12. Newman LA, Griffith KA, Jatoi I, Simon MS, Crowe JP, Colditz GA, *et al.* Meta-analysis of survival in African American and white American patients with breast cancer: Ethnicity compared with socioeconomic status. J Clin Oncol 2006;24:1342-9.
13. Gathani T, Ali R, Balkwill A, Green J, Reeves G, Beral V, *et al.* Ethnic differences in breast cancer incidence in England are due to differences in known risk factors for the disease: Prospective study. Br J Cancer 2014;110:224-9.
14. Cunningham R, Shaw C, Blakely T, Atkinson J, Sarfati D. Ethnic and socioeconomic trends in breast cancer incidence in New Zealand. BMC Cancer 2010;10:674.
15. Hussain SK, Altieri A, Sundquist J, Hemminki K. Influence of education level on breast cancer risk and survival in Sweden between 1990 and 2004. Int J Cancer 2008;122:165-9.
16. Herndon JE 2nd, Kornblith AB, Holland JC, Paskett ED. Effect of socioeconomic status as measured by education level on survival in breast cancer clinical trials. Psychooncology 2013;22:315-23.
17. Blakely T, Shaw C, Atkinson J, Cunningham R, Sarfati D. Social inequalities or inequities in cancer incidence? Repeated census-cancer cohort studies, New Zealand 1981-1986 to 2001-2004. Cancer Causes Control 2011;22:1307-18.
18. Hossain SZ, Robinson L, Clarke J. Breast cancer knowledge and participation in breast screening practices among Southeast Asian women living in Sydney. GSTF J Nurs Healthc 2016;3:82-92.
19. Glied S, Lleras-Muney A. Technological innovation and inequality in health. Demography 2008;45:741-61.
20. Griggs JJ, Culakova E, Sorbero ME, Poniewierski MS, Wolff DA, Crawford J, *et al.* Social and racial differences in selection of breast cancer adjuvant chemotherapy regimens. J Clin Oncol 2007;25:2522-7.
21. Kroenke CH, Kubzansky LD, Schernhammer ES, Holmes MD, Kawachi I. Social networks, social support, and survival after breast cancer diagnosis. J Clin Oncol 2006;24:1105-11.
22. Carmichael AR. Obesity as a risk factor for development and poor prognosis of breast cancer. BJOG 2006;113:1160-6.
23. Hauner D, Janni W, Rack B, Hauner H. The effect of overweight and nutrition on prognosis in breast cancer. Dtsch Arztebl Int 2011;108:795-801.
24. Beasley JM, Newcomb PA, Trentham-Dietz A, Hampton JM, Bersch AJ, Passarelli MN, *et al.* Post-diagnosis dietary factors and survival after invasive breast cancer. Breast Cancer Res Treat 2011;128:229-36.
25. Xing MY, Xu SZ, Shen P. Effect of low-fat diet on breast cancer survival: A meta-analysis. Asian Pac J Cancer Prev 2014;15:1141-4.
26. World Cancer Research Fund/American Institute for Cancer Research. Continuous Update Project Report Summary. Food, Nutrition, Physical Activity, and the Prevention of Breast Cancer. London: World Cancer Research Fund/American Institute for Cancer Research; 2010.
27. Narod SA. Tumour size predicts long-term survival among women with lymph node-positive breast cancer. Curr Oncol 2012;19:249-53.