

ORIGINAL RESEARCH

Usefulness of Gail Model Breast Cancer Risk Assessment Tool in Estimating the Risk for Development of Breast Cancer in Women of Kerala India

Paul Augustine,^a Regi Jose,^b Anoop Amrithlal,^c Zinia T Nujum,^d Arun Peter,^a Jeessa C Haran^b

a. Division of Surgical Oncology, Regional Cancer Center, Thiruvananthapuram, Kerala, India; b. Department of Community Medicine, Sree Gokulam Medical College & Research Foundation, Thiruvananthapuram, Kerala, India; c. Department of Community Medicine, Azeezia Institute of Medical Sciences, Kollam, Kerala, India; d. Department of Community Medicine, Government Medical College, Thiruvananthapuram, Kerala, India*

Corresponding Author: Dr. Regi Jose, Department of Community Medicine, Sree Gokulam Medical College & Research Foundation, Thiruvananthapuram, Kerala, India. Email: regipaul@gmail.com

Abstract

Published on 28th December, 2015

Breast cancer is the most common cancer among females in the state of Kerala, India. Gail *et al.* model is considered as a good tool to estimate woman's risk of developing breast cancer and are useful in targeting screening and prevention efforts. To determine the usefulness of Gail model breast cancer risk assessment tool in identifying women at high risk for breast cancer in Kerala. A case-control Study was conducted at Regional Cancer Center, Thiruvananthapuram, Kerala, India by including all breast cancer patients admitted for surgery from 1st of September 2003 to 31st December 2004 (Case n = 660 and controls n = 920). Participants were interviewed using a proforma. Gail's tool was used to calculate risk. The participants were grouped as a high, normal, or Low risk with regard to their risk in comparison with the general population score given in the calculator. Sensitivity and specificity of the model were found out. Unconditional logistic regression was used to estimate odds ratio 95% confidence intervals

(CIs) and for the final model. The mean score of cases was 0.872 (standard deviation [SD] - 0.460) and that of controls was 0.731 (SD - 0.403) (t-5.39 p-value 0.001). Overall sensitivity is 14.2 and specificity are 89.2. The major risk factors of breast cancer were age, irregular periods, previous history of breast biopsy, presence of first-degree relatives with breast cancer, history of abortion, absence of live birth, late age at first live birth, post-menopausal status, and absence of breast feeding. Breast feeding provides protection against breast cancer. Age of menarche was not found to be as a risk factor for breast cancer. A new model was also made using the identified risk factors. Gail model cannot be used to predict high-risk women in Kerala. A new model formulated based on the identified risk factors should be more useful in community-wide screening programs in Kerala.

Key Words: Breast Cancer, Gail Model, Validation, Risk Factors, India

Cite this article as: Augustine P, Jose R, Amrithlal A, Nujum ZT, Peter A, Haran JC. Usefulness of Gail Model Breast Cancer Risk Assessment Tool in Estimating the Risk for Development of Breast Cancer in Women of Kerala, India. Academic Medical Journal of India. 2015 Dec 28;3(4):117-122.

Introduction

Breast cancer is the most common malignancy among the women of Kerala,¹⁻³ yet one of the most treatable if detected early. Early detection and proper treatment are essential to obtain the desired results and longevity.

The Gail Model

The multivariate risk model developed by Gail *et al.*⁴ is widely used for quantifying the risk of breast cancer within a specified time in women at various ages and with certain risk factors. It was used in the National Surgical Adjuvant Breast and Bowel Project P-1 Study (more commonly known as the breast cancer prevention trial),⁵ and it accurately predicted the incidence of invasive breast cancer among women receiving placebo. The model has also been validated in two independent evaluations in separate populations.^{1,6} Only five factors were shown to be significant predictors of lifetime risk of breast cancer: (1) Current age, (2) age at menarche, (3) number of benign breast biopsies, (4) age at first live birth (FLB), and (5) family history of breast cancer in first-degree relatives.

Biopsies (incision, excision, or fine-needle aspirations, but not cyst aspirations) for benign breast disease are considered, and a biopsy showing atypical hyperplasia carries twice the risk of a biopsy showing no benign disease.

Alternatives to the Gail model^{7,8} allow consideration of the number of both first-degree and second-degree relatives with breast cancer as well as their ages at diagnosis. More complete risk assessment tables are available in the medical literature, and all permit calculation of a patient's lifetime probability of breast cancer.

Artists and scholars have studied the breast for centuries. It has been viewed as an object of beauty and a symbol of motherhood. Artists have captured the breast in the paint, marble, and bronze; However, we are yet to capture the reason why some women develop breast cancer and others do not. Who is at risk? Who is at high risk? And, can we prevent breast cancer in high-risk patients?

Breast cancer is the most common malignancy among the women of Kerala, and one of the most treatable if detected

*See End Note for complete author details

early. Early detection is essential to obtain the desired remission and longevity. Numerous studies have shown that advancing age, early age of menarche, age of FLB, family history of breast cancer, and number of previous breast biopsies were the major risk factors of breast cancer. Numerous mathematical models were also made using these risk factors to predict individual women's risk of developing breast cancer. Gail model is one such model formulated by Gail *et al.* of the National cancer institute Washington DC and was extensively used for prediction and identifying high-risk women for screening. There is a paucity of data regarding the risk factors in our society. A study of the risk factors of breast cancer and validation of Gail model would throw some light on this aspect, and we could use similar models to predict the high-risk women in our society and encourage them for screening. Risk factors, which are implicated in the etiology of breast cancer in a western setting, may not necessarily be true in our country. More epidemiological studies are required to elicit correlation, if any, in the Indian context.^{9,10}

Objectives

To determine the usefulness of Gail model breast cancer risk assessment tool in identifying women at high risk for getting breast cancer.

Methods

Study Design

Case-control study.

Study Setting

Hospital - Regional Cancer Center, Thiruvananthapuram, and Community - Thiruvananthapuram corporation area.

Study Period

22 Months (June 2003-March 2005).

Study Subjects

Case: Those newly diagnosed with breast cancer admitted for surgery during the study period. Control: Those who do not have breast cancer were selected from the hospital and the community. Hospital controls: Accompanying persons of patients (other than breast cancer) admitted to RCC, who did not have a history of breast cancer. Community controls: Women of Trivandrum corporation area, who did not have a history of breast cancer.

Exclusion Criteria

First-degree relatives of the cases under study and persons diagnosed with gynecological malignancy/contralateral breast cancer were excluded.

Sample Size

The sample size was fixed based on the results of the pilot study. The sample size needed for studying the risk factors was 486 based on odds ratio (OR). Gail model was only for the 35 years and above age group, for validating the 5-year prediction, women 40 years and above were needed. Hence, for validation, sample size was fixed for those above the age of 40 years based on the sensitivity (sensitivity 43.6%, precision 10% (0.0436) 95% confidence interval [CI]; $N = [(Z_{1-\alpha/2})^2 P(1-P)]/d^2$ $n = 497$). Finally, giving due consideration to all factors, it was decided that all patients should be included until 500 cases of 40 years, or older was obtained; and a total sample of 660 cases.

Sampling

Case: All incident cases admitted for surgery during the study period until the required number is obtained were included in the study. Controls: Hospital - Bystanders of other cancer patients admitted for surgery in the same ward next to the participant with breast cancer. Community - 30 women identified from each of the 15 clusters selected randomly from the 81 wards of Trivandrum corporation area.

Methods

Gail's score is used to calculate risks of females above the age of 35 and predicts the 5-year and lifetime risk of developing breast cancer. The tool also provides the risk for the general population to develop breast cancer and classify them as high or normal risk. This was designed using a cohort study. The present study - a case-control study, is planned to validate this tool for predicting 5-year risk for developing breast cancer in the study group. The study group would be interviewed with the help of a proforma after obtaining informed consent.

The subject is asked to recollect the five risk factors, which were used in Gail score, at a date 5 years before the development of breast cancer. The factors of interest are constant and have a personal concern, and hence, the problem of recall bias does not arise. The data are then entered in the Gail's calculator; computer software, and the score is calculated. The mean scores of both cases and controls of each age group (2 years groups) were found out. Then, the results of both groups were compared to see any association with the Gail's score. The participants were also grouped as a high, normal, or low risk with regard to their risk in comparison with the general population score given in the calculator. Data regarding risk factors were also obtained during the interview and were analyzed to find out the crude OR.

Statistical Analysis

Univariate analysis

Data on continuous variables were summarized as means and standard deviation (SD). Unpaired t-test was used to test for

statistical significance. Categorical variables were summarized as proportions, and the chi-square test was used in such situations. Crude estimates of risk were made for both continuous and categorical variables using Logistic regression (Binary logistic-Enter method) and expressed as crude OR with 95% CI.

Multivariate Analysis

Independent variables were defined as those with a significant OR during the crude analysis. The dependent variable was case-control status. Logistic regression modeling was done using SPSS software. Factors that continued to have significant OR after adjustment (adjusted OR within 95% CI) were included in the final model. The results include; measures of risk presented as adjusted OR with 95% CI, the final model, and tests for its validity.

Validation of Gail Model

The usefulness of Gail model was studied using sensitivity, specificity, positive and negative predictive values, likelihood ratios (LRs), and Youden's J test.

Ethical Considerations

The Institutional Review Board Scientific Review Committee Regional Cancer Center (RCC) has sanctioned the research study at its meeting on the August 30th, 2003. Obtained clearance from the Ethical Committee of Medical College, Trivandrum and RCC, Trivandrum.

Results and Discussion

Results of Univariate Analysis

General characteristics

Age

The study participants comprised cases with ages ranging from 20 to 85 years and controls with ages ranging from 19 to 77 years. The mean age of cases was 47.79 (SD - 11.1) and that of control was 44.77 (SD - 10.93). Means were tested using t-test and were found to be significant (Significance - 0.001). The median age of cases was 47 years and controls 43 years (**Figure 1**).

Crude OR was tested using logistic regression = 1.025 (95% CI 1.016-1.034), which was found to be significant. For each year of age advances, the risk increases by 2.5%.

In India, breast cancer occurs a decade earlier than western women—the mean age of occurrence is about 42 in India compared to 53 in the white women.

In the United States, the risk of breast cancer is higher in middle-aged and elderly women than in young women.¹¹

Age of menarche

In the present study, the mean age of menarche for cases was 14.25 ± 1.51 and for controls was 14.329 ± 1.56 was

not statistically significant (t-0.93 P = -0.402). The crude OR - 0.970; (95% CI - 0.908-1.035) was also not significant.

Age of onset of menarche was classified into two groups with 12 as the cutoff age and tested for significance, which showed no significant difference in the distribution among cases and controls (Chi-square - 1.218; df - 1; Significance - 0.270). Crude OR - 1.209 (95% CI - 0.863-1.695) (**Table 1**).

This finding is against the common finding of early age at menarche as a risk factor of breast cancer.¹² However, studies from India Reddy *et al.* show that age at menarche appears to have no association with the risk of breast cancer. Moreover, Gajalakshmy *et al.* analyzed the risk factors for breast cancer separately in the pre- and the post-menopausal groups. In neither group was there a significant association between age at menarche and breast cancer risk. According to Helmrich *et al.*, late age at menarche was associated with a lower risk among pre-menopausal women but not among post-menopausal women.

Age of FLB

Late age of FLB is another important risk factor with relative risk ranging from 1.9 to 3.5 for those who have their first child after 30 years compared to those who have their first child before 20 years.

In the present study, of the 1435 women with at least one child; the mean age of FLB among cases was 23.99 ± 4.36 and

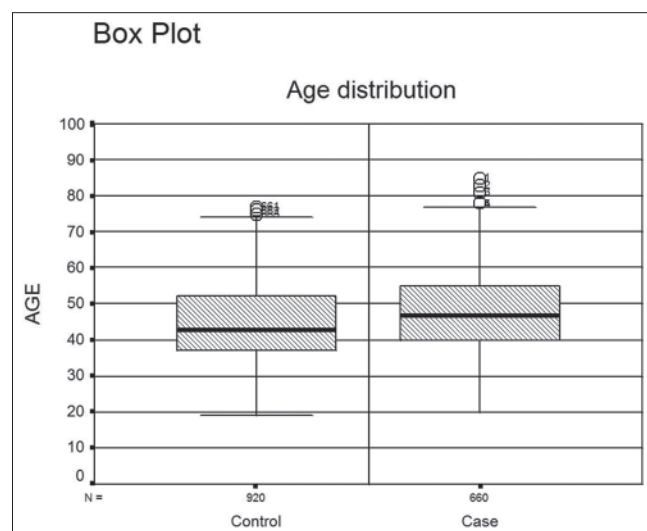


Figure 1: Age distribution of study participants

Table 1: Age of menarche in two groups

Age of menarche	Control (%)	Case (%)	Total (%)
>12 years	839 (91.2)	591 (89.5)	1430 (90.5)
≤12 years	81 (8.8)	69 (10.5)	150 (9.5)
Total	920 (100.0)	660 (100.0)	1580 (100.0)

Pearson Chi-square - 1.218; df - 1; P value=0.27; Crude OR - 1.209; 95% CI - 0.863-1.695. OR: Odds ratio, CI: Confidence interval

of controls was 22.28 ± 3.78 . Means were tested for significance using t-test and were found to be significant ($t = 7.870$; p -value 0.001). The Crude OR was also tested which was found to be significant (Crude OR - 1.109; [95% CI - 1.080-1.140]) suggesting that the penalty for waiting each year for having the first child will increase the chance of getting breast cancer by 10%.

A study from Jaipur showed most of the women (68.50%) had their first live child between 18 and 25 years and 13.50% had below 18 years.¹⁰ Late age at FLB was associated with increased risk of breast cancer among both pre- and post-menopausal women (Gao *et al.*).

Categories of Age of FLB

Age of FLB was coded into four categories to compare the odds across the age groups, and nulliparous women were included in the last category of the age of FLB after 30 years. **Table 2** depicts the odds ratio with CI and shows that those with the age of FLB 30 years or more and Nullipara has a 5-fold risk for developing breast cancer compared to women with the first child before 20 years.

This agrees with Rao *et al.* and many others, that women who delayed their first childbirth were at elevated risk of developing breast cancer. According to Tavani *et al.*, breast cancer was directly associated with age at first birth (OR - 5.31) among women aged 30 years or more at first birth compared with those aged <20.

Number of Previous Breast Biopsies

Participants were asked about the history of breast biopsies prior to the diagnosis. This study also says that the OR is significantly higher for those who had a previous breast biopsy and the risk increases 11-fold if the person has more than one breast biopsy.

History of Biopsy

The frequencies in **Table 3** may give a false representation with regard to OR, and so it was further grouped into presence or absence of the previous biopsy. **Table 4** depicts the frequencies across cases and controls and is significant. The crude OR 3.663 (95% CI - 2.183-6.141) is significant which

says those who had a history of previous breast biopsy has a 3.6-fold more chance for developing breast cancer compared with those who never had a biopsy.

These findings agree with other studies which showed that prior breast biopsy is associated with increased risk of breast cancer.

Number of First-degree Relatives (FRL) With Breast Cancer

Number of FRL with breast cancer was one of the major risk factors for breast cancer and it is included in risk prediction using Gail model⁴ and Claus⁸ model.

In this study, 4.7% of cases have one first-degree relative, and one control had two relatives. So, **Table 5** was reclassified based on the presence or absence of the risk factor. **Table 5** shows 4.7% of cases had a positive family history of breast cancer. Those who have a first-degree relative with breast cancer has a 2-fold risk 2.218 (95% CI - 1.253-3.927) of developing breast cancer in later life.

This agrees with many studies, which showed the presence of a first-degree relative with breast cancer as risk factor of breast cancer.^{13,14} However, according to Ramachandra *et al.*, family history was rare in those with breast cancer and not a risk factor of breast cancer.

Validation of Gail Model Breast Cancer Risk Assessment Tool

Participants for Gail model validation included those 40 years or above. Gail scores of all participants 40 years or more were calculated based on their risk factors 5 years back. There were 500 cases and 600 controls. The mean score of cases was 0.872 (SD - 0.460) and that of controls was 0.731 (SD - 0.403). The t-test was performed as the test of significance and is found to be significant ($t = 5.392$; Significance - 0.000).

Classification of Study Participants

High risk - those participants who have a score greater than the general population score given in Gail model for comparison.
Equal risk - those participants who have a score equal to the general population score given in Gail model for comparison.

Table 2: Categories of age at first live birth

Age category (years)	Control (%)	Case (%)	Total (%)	Exp (B)	95.0% CI for experiment (B)	
					Lower	Upper
<20	216 (23.5)	74 (11.2)	290 (18.4)	1.00	Nil	Nil
20-24	445 (48.4)	257 (38.9)	702 (44.4)	1.686	1.242	2.288
25-29	176 (19.1)	177 (26.8)	353 (22.3)	2.936	2.097	4.110
≥30 and Nullipara	83 (9.0)	152 (23.0)	235 (14.9)	5.345	3.671	7.784
Total	920 (100.0)	660 (100.0)	1580 (100.0)	Nil		

Pearson Chi-square - 100.066; df - 3; Significance - 0.000, CI: Confidence interval

Table 3: Number of previous breast biopsies

Number of biopsies (%)	Control (%)	Case (%)	Total (%)	Exp (B)	95.0% CI for Experiment (B)	
					Lower	Upper
No biopsy	899 (97.7)	608 (92.1)	1507 (95.4)	1.00	Nil	Nil
1 biopsy	20 (2.2)	44 (6.7)	64 (4.1)	3.253	1.899	5.574
>1 biopsy	1 (0.1)	8 (1.2)	9 (0.6)	11.797	1.475	94.355
Total	920 (100.0)	660 (100.0)	1580 (100.0)		Nil	

Chi-square=28.627; df - 2; p-value 0.001, CI: Confidence interval

Table 4: History of previous biopsy

History of biopsy	Control (%)	Case (%)	Total (%)
No	899 (97.7)	608 (92.1)	1507 (95.4)
Yes	21 (2.3)	52 (7.9)	73 (4.6)
Total	920 (100.0)	660 (100.0)	1580 (100.0)

Pearson Chi-square - 27.311; df - 1; p value - 0.001. Crude odds ratio=3.663 (95% CI - 2.183-6.141)

Table 5: FRL with breast cancer

FRL breast cancer	Control (%)	Case (%)	Total (%)
No	900 (97.8)	629 (95.3)	1529 (96.8)
Yes	20 (2.2)	31 (4.7)	51 (3.2)
Total	920 (100.0)	660 (100.0)	1580 (100.0)

Pearson Chi-square - 7.832; df - 1; p-value 0.005; Crude odds ratio=2.218 (95% CI=1.253-3.927). CI: Confidence interval, FRL: First-degree relatives

Low risks - those participants who have a score less than the general population score given in Gail model for comparison.

Table 6 while considering all participants, there is a significant (p-value 0.01) difference in the distribution of risk groups among cases and controls. The distribution across different age strata was also tested with Chi-square tests, which showed a significant difference (p-value 0.001) only for those 50 years or more.

Validation of Gail Model

The risk groups are further reclassified to study the sensitivity and specificity.

Table 7 shows true high risk - those with scores more than the comparison group.

Normal risk - those with scores equal to or less than the comparison group.

Table 7 shows that when considering all the participants the overall sensitivity is 14.2 and specificity is 89.2. The positive and negative predictive value is 52.2 and 55.4, respectively.

The distribution of the predicted risk groups was not significant (p-value 0.09) among cases and controls when considering all together. Across two age groups, Gail model was more sensitive (17.2) in age <50, compared to those 50 years or

Table 6: Gail risk comparison

Participants	Risk	Case	Control	Total	Chi-square	Significance
All	High	71	65	136	9.2	0.01
	Equal	101	91	192		
	Low	328	444	772		
	Total	500	600	1100		
<50	High	41	46	87	3.24	0.197
	Equal	79	83	162		
	Low	118	173	291		
	Total	238	302	540		
≥50	High	30	19	49	14.48	0.001
	Equal	22	8	30		
	Low	210	271	481		
	Total	298	262	560		

above (sensitive - 11.4), whereas specificity (93.6) was more in those 50 years or older.

The above finding is not significant (p-value 0.531) and is contradictory to most validation studies, which says that Gail model is more sensitive among older women.⁶

Predictive accuracy (Percentage agreement) was similar in all groups (0.55).

Conclusion

The purpose of the study was to determine the OR of the known risk factors and to evaluate the ability of Gail model to accurately predict individual breast cancer risk, using a population independent of the one from which the model was derived. The Gail model was found to be less sensitive; hence, a search for other possible risk factors was done to develop a tool that could be of use in the community of the study group.

The present study found that the major risk factors among the study population were – increasing age (OR = 1.047), previous breast biopsy (OR = 4.170), first-degree relatives with breast cancer (OR = 2.137), late age at FLB (age ≥30: OR = 3.705), not breastfeeding (OR = 1.969), irregular menstrual cycles (OR = 1.536), post-menopausal status (OR = 1.427), and past history of abortion (OR = 1.320).

Table 7: Tests for validation of Gail model

True high	Predicted risk	Case	Control	Total	Significance	Sensitivity	Specificity	PPV	NPV
All (%)	High	71 (14.2)	65 (10.8)	136	0.09	14.2	89.2	52.2	55.4
	Normal	429 (85.8)	535 (89.2)	966					
	Total	500 (100)	600	1100					
Age < 50	High	41	46	87	0.53	17.2	84.7	47.1	55.3
	Normal	197	256	453					
	Total	238	302	540					
Age ≥ 50	High	30	19	49	0.03	11.4	93.6	61.2	54.6
	Normal	232	279	511					
	Total	262	298	560					

PPV: Positive predictive value, NPV: Negative predictive value

These factors were considered, and a logistic regression model was formulated to calculate the probability of the participant to become a case of breast cancer.

Assessment of a woman's risk of breast cancer can be used for counseling and decision making about clinical management of risk. Physicians should give patients a clear and positive message regarding risk management and should emphasize that risk calculations are estimates only of the probability of having breast cancer, not the risk of dying of the disease. To conclude Gail's tool though useful in the US, may not be useful to identify high-risk women in our society.

End Note

Author Information

1. Dr. Paul Augustine, Division of Surgical Oncology, Regional Cancer Center, Thiruvananthapuram, Kerala, India
2. Dr. Regi Jose, Department of Community Medicine, Sree Gokulam Medical College and Research Foundation, Thiruvananthapuram, Kerala, India
3. Dr. Anoop Amrithlal, Department of Community Medicine, Azeezia Institute of Medical Sciences, Kollam, Thiruvananthapuram, Kerala, India
4. Dr. Zinia T Nujum, Department of Community Medicine, Government Medical College, Thiruvananthapuram, Kerala, India
5. Dr. Arun Peter, Division of Surgical Oncology, Regional Cancer Center, Thiruvananthapuram, Kerala, India
6. Dr. Jeesha C Haran, Department of Community Medicine, Sree Gokulam Medical College & Research Foundation, Thiruvananthapuram, Kerala, India

Conflict of Interest

None declared.

References

1. Spiegelman D, Colditz GA, Hunter D, Hertzmark E. Validation of the Gail *et al.* model for predicting individual breast cancer risk. *J Natl Cancer Inst* 1994;86:600-7.
2. Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, *et al.* GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC Cancer Base No. 11, Lyon, France: International Agency for Research on Cancer; 2013.
3. RCC. Hospital Cancer Registry 2013 Regional Cancer Centre, Thiruvananthapuram. Trivandrum: Regional Cancer Center, Trivandrum; 2014. Available from: <http://www.rcctvm.org>. [Last cited on 2014 Feb 04].
4. Gail MH, Brinton LA, Byar DP, Corle DK, Green SB, Schairer C, *et al.* Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J Natl Cancer Inst* 1989;81:1879-86.
5. Fisher B, Costantino JP, Wickerham DL, Redmond CK, Kavanah M, Cronin WM, *et al.* Tamoxifen for prevention of breast cancer: Report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. *J Natl Cancer Inst* 1998;90:1371-88.
6. Bondy ML, Lustbader ED, Halabi S, Ross E, Vogel VG. Validation of a breast cancer risk assessment model in women with a positive family history. *J Natl Cancer Inst* 1994;86:620-5.
7. Claus EB, Risch N, Thompson WD. Autosomal dominant inheritance of early-onset breast cancer. Implications for risk prediction. *Cancer* 1994;73:643-51.
8. Parmigiani G, Berry D, Aguilar O. Determining carrier probabilities for breast cancer-susceptibility genes BRCA1 and BRCA2. *Am J Hum Genet* 1998;62:145-58.
9. Gupta P, Sharma RG, Verma M. Review of breast cancer cases in Jaipur region. *J Indian Med Assoc* 2002;100:282-3, 286-7.
10. Vidyashankar C. Reuters health 'Western' risk factors for breast cancer do not seem to hold in India.' *J Indian Med Assoc* 2002;100:282-7.
11. Mayberry RM. Age-specific patterns of association between breast cancer and risk factors in black women, ages 20 to 39 and 40 to 54. *Ann Epidemiol* 1994;4:205-13.
12. Clemons M, Goss P. Estrogen and the risk of breast cancer. *N Engl J Med* 2001;344:276-85.
13. Hirose K, Tajima K, Hamajima N, Inoue M, Takezaki T, Kuroishi T, *et al.* A large-scale, hospital-based case-control study of risk factors of breast cancer according to menopausal status. *Jpn J Cancer Res* 1995;86:146-54.
14. Augustine P, Jose R, Peter A, Lal AA, Prabhakar J, Sreedharan J, *et al.* Risk Factors of Breast Cancer in Kerala, India - A Case Control Study. *Acad Med J India* 2014;2:7-13.