Discriminatory Accuracy of the Gail Model for Breast Cancer Risk Assessment among Iranian Women

Sahar ROSTAMI 1,2, Ali RAFEI 2, Maryam DAMGHANIAN 3, Zohreh KHAKBAZAN 3, Farzad MALEKI 2,4, *Kazem ZENDEHDEL 2,5,6

1. Department of Reproductive Health and Midwifery, School of Nursing and Midwifery, Tehran University of Medical Sciences, Tehran, Iran
2. Cancer Research Center, Cancer Institute of Iran, Tehran University of Medical Sciences, Tehran, Iran
3. Nursing and Midwifery Care Research Center, Tehran University of Medical Sciences, Tehran, Iran
4. Social Determinants of Health Research Center, Urmia University of Medical Sciences, Urmia, Iran
5. Cancer Biology Research Center, Cancer Institute of Iran, Tehran University of Medical Sciences, Tehran, Iran
6. Breast Disease Research Center, Cancer Institute of Iran, Tehran University of Medical Sciences, Tehran, Iran

*Corresponding Author: Email: kzendeh@tums.ac.ir

(Received 14 May 2019; accepted 17 Jul 2019)

Abstract
Background: The Gail model is the most well-known tool for breast cancer risk assessment worldwide. Although it was validated in various Western populations, inconsistent results were reported from Asian populations. We used data from a large case-control study and evaluated the discriminatory accuracy of the Gail model for breast cancer risk assessment among the Iranian female population.

Methods: We used data from 942 breast cancer patients and 975 healthy controls at the Cancer Institute of Iran, Tehran, Iran, in 2016. We refitted the Gail model to our case-control data (the IR-Gail model). We compared the discriminatory power of the IR-Gail with the original Gail model, using ROC curve analyses and estimation of the area under the ROC curve (AUC).

Results: Except for the history of biopsies that showed an extremely high relative risk (OR=9.1), the observed ORs were similar to the estimates observed in Gail’s study. Incidence rates of breast cancer were extremely lower in Iran than in the USA, leading to a lower average absolute risk among the Iranian population (2.78, ±SD 2.45). The AUC was significantly improved after refitting the model, but it remained modest (0.636 vs. 0.627, ΔAUC = 0.009, bootstrapped P=0.008). We reported that the cut-point of 1.67 suggested in the Gail study did not discriminate between breast cancer patients and controls among the Iranian female population.

Conclusion: Although the coefficients from the local study improved the discriminatory accuracy of the model, it remained modest. Cohort studies are warranted to evaluate the validity of the model for Iranian women.

Keywords: Breast neoplasms; Risk assessment; Models, Statistical; Logistic models

Introduction

With an estimated 2 million cancer cases diagnosed in 2018 (24% of all cancers), breast cancer is the most common malignancy among women worldwide (1). Incidence and mortality rates of breast cancer are growing in low and middle-income countries (LMICs), including Iran (2-4). Because of the younger age structure of the population in the LMICs, the proportion of premenopausal breast cancer is significantly higher in these countries compared to the high-income countries.
(HICs) (5). Therefore, there are debates to start prevention programs at an earlier age in the LMICs (6). Mammography screening is the most promising tool for early detection of breast cancer (2), and it has led to a decrease in the mortality rate of breast cancer worldwide (7-9). However, due to a higher rate of false-positive reports in young women, it is not an appropriate screening test for this age group (10). Moreover, mammography screening is not cost-effective in Iran and other LMICs (11, 12). Thus, risk assessment tools can help to define high-risk groups and design personalized screening programs in these countries (13-15).

The Gail model was developed in 1989 has been used widely to predict the risk of breast cancer based on the major risk factors (16). Later, by incorporating the age-specific breast cancer rates from Surveillance, Epidemiology, and End Results Program (SEER), it was modified to "Gail model 2" to estimate absolute risk for developing invasive breast cancer (17). The original model was developed in the USA, where the association of the risk factors and the incidence and mortality rates of breast cancer are different from other countries (15), especially from the Asian population and other LMICs (18-21). Several studies evaluated the validity and accuracy of the Gail model in different countries and specific ethnic populations in the USA (22-24). Gail model is valid for estimating the 5-year and lifetime risk of breast cancer in different populations, especially in the USA and Europe (23). Therefore, the model seems to be a useful tool for risk assessment in the clinics and breast cancer prevention programs such as tamoxifen prescription (25).

Despite the validity of the Gail model for risk assessment, the discriminatory power (the area under the ROC curve or AUC statistic) of the model has been modest in almost all studies. The summary of the AUC in a meta-analysis of 29 studies was 0.60 (95% CI 0.58-0.62), indicating that the model is not working accurately in identifying the high-risk groups for population-based interventions (23). Therefore, several research groups tried to add other risk factors, including clinical and genetic data, to improve the sensitivity and specificity of the model and optimize its discriminatory power for risk assessment in different populations (24). The most promising results were reported in the USA (26) and Sweden (27,28), in which the AUC increased up to about 0.70.

A few studies evaluated the performance of the Gail model among Iranian women (29-35). However, they used a small sample size and did not use local estimates for breast cancer incidence and mortality rates. We used local estimates and utilized ORs from a large case-control study in Iran and studied the discriminatory accuracy of the Gail model for breast cancer risk assessment among the Iranian female population.

Materials and Methods

Study design

We used data from a hospital-based, case-control study conducted at the Cancer Institute of Iran, Tehran, Iran, in 2016. The methods of recruitment of the cases and controls and study design have been described in detail elsewhere (36). In brief, cases were 942 incident patients, and controls were 975 healthy visitors who were frequency-matched for age in five-year categories and residential place.

Statistical analysis

In the current study, similar to Gail et al. (16), we approximated relative risks (RRs) by estimating odds ratios (ORs) obtained from a logistic regression model that included age at menarche (AGEMEN) [coded as ≥14 (0), 12-13 (1), or <12 (2)], number of previous breast biopsies (NBIOPS) [coded as 0 (0), 1 (1), or ≥2 (2)], age at first live birth (AGEFLB) [coded as <20 (0), 20-24 (1), 25-29 or nulliparous (2), or ≥30 (3)], and number of first-degree relatives (mother or sisters) with breast cancer (NUMREL) [coded as 0 (0), 1 (1), or ≥2 (2)], as well as AGECAT [categorized as <50 or ≥50] and interactions between AGECAT and NBIOPS and between AGEFLB and NUMREL.
Variables were coded according to the categories provided in the original Gail model (16). We did not collect information on the presence of atypical hyperplasia on previous biopsies in this study. Absolute risk is the probability that a healthy subject free of cancer at age α will develop breast cancer in a subsequent age interval. Suppose [α, α + Δ] is the time interval of interest. Then, the absolute risk is given by:

\[ P(\alpha, \Delta, r(t)) = \int_{\alpha}^{\alpha+\Delta} h_1(t) r(t) \exp\{-\int_{\alpha}^{t} h_1(u) r(u) \, du\} \frac{S_2(t)}{S_2(\alpha)} \, dt \]

Where \( S_2(t) = \exp\{-\int_{\alpha}^{t} h_1(u) r(u) \, du\} \) is the probability of surviving competing risks (causes of death other than breast cancer) up to age t, the age at which a woman will survive without developing breast cancer. In this equation, the term \( S_2(t)/S_2(\alpha) \) corresponds to the conditional probability of surviving competing risks from age α to t. The exponential term corresponds to surviving without breast cancer from age α to age t. There is an instantaneous probability \( h_1(t) r(t) dt \) of developing breast cancer at age t. The baseline hazard, \( h_1(t) \), is estimated by multiplying age-specific breast cancer incidence rates \( h_2(t) \), by a conversion factor equal to one minus the population attributable risk (1-AR). The age-specific risk of dying from causes other than breast cancer is represented by \( h_2(t) \) and is assumed to be the same for all individuals. We assumed \( h_1, h_2, \) and \( r \) to be constant within 5-year intervals.

We combined the case-control data with the estimated age-specific breast cancer incidence rate (ASR) from the population-based cancer registry and the competing mortality rates from the national report on all-cause mortality among Iranian women. We converted RRs to 5-year and lifetime absolute risks using baseline age-specific breast cancer hazard rate and competing mortality rate. We assessed the performance of the modified model (the IR-Gail model) and studied the discriminatory power using ROC curve analyses and estimation of AUC. Goodness-of-fit was assessed using the Hosmer-Lemeshow test (P>0.05). We applied bootstrap resampling to calculate confidence intervals (CIs) for the AUC and P-values for differences in AUC. Considering three risk categories, we used a reclassification chart to assess the model regarding the assignment of women to low, intermediate, and high-risk categories, based on 5-year absolute risk estimates. Three cut-off values were determined by the first, second, and third quartiles of predicted risk by the Gail model. We used R statistical software (version 3.4.1) and utilized BCRA and ROCR packages for statistical analysis.

The study was approved by the Research Ethics Committee of TUMS (code: IR.TUMS.FNM.REC.1395.618).

**Results**

We used data from 942 breast cancer patients and 975 controls in this study for breast cancer risk assessment (Table 1). We found positive associations between breast cancer and different risk factors, including the previous breast biopsies (OR=9.12), age at first live birth (OR=1.5), positive family history of breast cancer (OR=2.52), and age (OR=1.25) (Table 2). The interaction between "number of biopsies" and "age" was statistically significant (P=0.003). ORs reported in the Gail study were almost similar to this study, except for the number of the previous biopsies, of which the OR was considerably higher (OR=9.12) than the estimate provided in Gail's (OR= 1.7).

Age-specific incidence rates of breast cancer were considerably lower than the rates used in the Gail study. The peak incidence rate was 110 per 100,000 in age 50-54 yr in Iranian women, while it was 443 per 100,000 in age 75-79 yr among the US. Despite a similar trend, the competing mortality rate was lower among Iranian women after the age of 85 years (Fig. 1). The average 5-year risks of invasive breast cancer in cases and controls were 0.70 (±1.73) and 0.37 (±1.49), respectively.
Table 1: Distribution of risk factors of the Gail model among cases and controls

<table>
<thead>
<tr>
<th>Predictive Factors (code NO.)</th>
<th>Cases (n=942)</th>
<th>Controls (n=975)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>NUMBIOP^1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 (0)</td>
<td>869</td>
<td>92.2</td>
</tr>
<tr>
<td>1 (1)</td>
<td>67</td>
<td>7.1</td>
</tr>
<tr>
<td>≥2 (2)</td>
<td>6</td>
<td>0.6</td>
</tr>
<tr>
<td>AGEFLB^2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;20 (0)</td>
<td>424</td>
<td>45</td>
</tr>
<tr>
<td>20- 24 (1)</td>
<td>281</td>
<td>29.8</td>
</tr>
<tr>
<td>25- 29 or nulliparous (2)</td>
<td>149</td>
<td>15.8</td>
</tr>
<tr>
<td>≥30 (3)</td>
<td>88</td>
<td>9.3</td>
</tr>
<tr>
<td>AGEMEN^3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥14 (0)</td>
<td>431</td>
<td>45.8</td>
</tr>
<tr>
<td>12- 13 (1)</td>
<td>377</td>
<td>40</td>
</tr>
<tr>
<td>&lt;12 (2)</td>
<td>134</td>
<td>14.2</td>
</tr>
<tr>
<td>NUMREL^4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 (0)</td>
<td>870</td>
<td>92.4</td>
</tr>
<tr>
<td>1 (1)</td>
<td>58</td>
<td>6.2</td>
</tr>
<tr>
<td>≥2 (2)</td>
<td>14</td>
<td>1.5</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;50 (0)</td>
<td>579</td>
<td>61.5</td>
</tr>
<tr>
<td>≥50 (1)</td>
<td>363</td>
<td>38.5</td>
</tr>
</tbody>
</table>

1 Number of previous breast biopsies
2 Age at first live birth
3 Age at menarche
4 Number of first-degree relatives (mother or sisters) with breast cancer

Table 2: Comparison of the coefficient estimates of logistic regression and relative risks between our study and Gail's model

<table>
<thead>
<tr>
<th>Predictive Factors</th>
<th>Iran</th>
<th>OR</th>
<th>P-value</th>
<th>Gail</th>
<th>OR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.57126</td>
<td>0.56</td>
<td>&lt;0.0001</td>
<td>-0.3286</td>
<td>0.47</td>
</tr>
<tr>
<td>AGEMEN</td>
<td>0.07689</td>
<td>1.08</td>
<td>0.2676</td>
<td>0.09401</td>
<td>1.1</td>
</tr>
<tr>
<td>NBIOPS</td>
<td>2.21061</td>
<td>9.12</td>
<td>&lt;0.0001</td>
<td>0.52926</td>
<td>1.7</td>
</tr>
<tr>
<td>AGEFLB</td>
<td>0.40709</td>
<td>1.5</td>
<td>&lt;0.0001</td>
<td>0.21863</td>
<td>1.24</td>
</tr>
<tr>
<td>NUMREL</td>
<td>0.92584</td>
<td>2.52</td>
<td>0.0002</td>
<td>0.95830</td>
<td>2.61</td>
</tr>
<tr>
<td>AGECAT</td>
<td>0.22382</td>
<td>1.25</td>
<td>0.0247</td>
<td>0.01081</td>
<td>1.01</td>
</tr>
<tr>
<td>NBIOPS*AGECAT</td>
<td>-1.65308</td>
<td>0.19</td>
<td>0.0032</td>
<td>0.28804</td>
<td>1.33</td>
</tr>
<tr>
<td>AGEFLB*NUMREL</td>
<td>-0.11903</td>
<td>0.89</td>
<td>0.6083</td>
<td>0.19081</td>
<td>1.21</td>
</tr>
</tbody>
</table>

The mean lifetime risk was 3.92 (±4.96) in cases and 2.78 (±2.45) in the control group (Table 3). We observed an AUC of 0.63 (95% CI=0.61-0.66) for the IR-Gail, compared with an AUC of 0.62 (95% CI=0.60-0.65) for the original Gail model. The difference in AUCs was statistically significant (ΔAUC = 0.009, bootstrapped P=0.008) (Fig. 2). The Hosmer-Lemeshow test revealed that the IR-Gail model fitted the data well (P=0.525).

Available at: http://ijph.tums.ac.ir
Fig. 1: Comparison of the age-specific incidence rates (ASRs) and mortality rates of breast cancer in Iran and the USA population.

Fig. 2: ROC analysis of the discriminatory accuracy of the Gail model among Iranian women.

Table 3: Descriptive statistics of age, and 5-year and lifetime risk of breast cancer in patients (n=942) and control (n=975).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cases</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (yr), Mean (SD)</td>
<td>47 (10.1)</td>
<td>45.1 (10.8)</td>
</tr>
<tr>
<td>5-year risk (%), Mean, SD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (±SD)</td>
<td>0.70 (1.73)</td>
<td>0.37 (1.49)</td>
</tr>
<tr>
<td>Median</td>
<td>0.38</td>
<td>0.31</td>
</tr>
<tr>
<td>N (%), higher than 1.67</td>
<td>57 (4%)</td>
<td>7 (1%)</td>
</tr>
<tr>
<td>N (%), lower than 1.67</td>
<td>885 (94%)</td>
<td>960 (99%)</td>
</tr>
<tr>
<td>Lifetime risk (%), Mean, SD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (±SD)</td>
<td>3.92 (4.96)</td>
<td>2.78 (2.45)</td>
</tr>
<tr>
<td>Median</td>
<td>2.64</td>
<td>2.30</td>
</tr>
</tbody>
</table>

Available at: http://ijph.tums.ac.ir
Discussion

We used data from a large case-control study and evaluated the discriminatory accuracy of the Gail model for breast cancer risk assessment. Previous breast biopsies had a strong association with the risk of breast cancer (OR=9.12) compared to the estimates provided in Gail's study (16), which could be due to the lack of a population-based screening program in Iran (3, 37). OR of other risk factors were almost similar to the reports provided by other countries, including the results of the Gail study. Although recalculating coefficients improved the predictive accuracy of the model significantly (bootstrapped P=0.008), similar to those from studies conducted in other countries, the AUC was modest, and it did not reach an appropriate threshold for classification of the female population to the high- and low-risk group in the population level.

A particular strength of this large, well-designed case-control study was the statistical power and precise estimation of relevant risk factors necessary for the Gail model. Patients from all over the country are visiting the Cancer Institute of Iran (37), indicating that data from this study could be a reasonable cross-representative sample from all over the country and generalizer the results to Iran. We used national estimates for the age-specific incidence rate of breast cancer and all-cause mortality among the Iranian population. However, this study faced some limitations. Our data totally consisted of the Iranian female population. As none of the race/ethnicity categories within the Gail model mapped to the ethnic profile of Asian women, we considered race status as unknown due to homogeneity. Different ethnic groups live in Iran, including Farses, Turks, Kurds, Arabs, etc. and their risk may vary. Unfortunately, we did not have data on the variation of cancer risk by ethnic groups to incorporate in the model. It remains a priority for the future when such data is available. Moreover, we lacked data on the history of atypical hyperplasia status for women with prior breast biopsies and considered them unknowns in the analysis. However, due to the lack of a comprehensive breast cancer screening program in Iran and a reasonably low incidence of atypical hyperplasia (3), this limitation is unlikely to affect the results of this study. Since opportunistic screening is increasing in Iran, we suggest reviewing pathology reports, including this variable and other screening related data such as breast density in future studies. The American Society of Clinical Oncology (ASCO) guideline recommended that women with a 5-year projected breast cancer risk of 1.67% or greater, which is that of an average 60-year-old woman, may benefit from chemopreventive by tamoxifen (38). This cut point was first used in the Breast Cancer Prevention Trial (BCPT), a trial that reported an almost 50% reduction in breast cancer risk among women given tamoxifen (39). However, the Gail model uses data from white American women. Using breast cancer incidence and mortality rates in each population may improve the calibration of the Gail model (40, 41).

Several studies evaluated the discriminatory accuracy of the Gail model. Based on a recent meta-analysis (23), the summary of AUC obtained from these studies was about 0.60, indicating that the power of the Gail model for discrimination between breast cancer patients and the general healthy population is relatively low. The AUC in this study was slightly higher (0.63), indicating that the discriminatory role of the Gail model is similar to other populations and the Gail model is applicable for Iranian women. Unfortunately, we could not evaluate the calibration of the model and the accuracy of the cut points in a cohort study. A recent meta-analysis based on 29 studies showed that the Gail model slightly overestimates the risk of breast cancer (23). Summary estimates of the expected to the observed ratio (E/O) were 1.16 (95% CI: 1.05-1.30). The summary E/O ratio was 1.02 (95% CI: 0.93-1.12) in the US and 1.05 (95% CI: 0.68-1.63) in Europe. However, the summary of the E/O ratio was considerably higher in Asian countries (E/O: 1.98, 95% CI: 1.58-2.48). No study has so far evaluated the validity of the Gail model in Iran and other Western Asian and African countries. Cohort studies are warranted to

Available at:  http://ijph.tums.ac.ir
verify the validity of the Gail model in Iran and other countries in this region.

Several research groups tried to include additional behavioral, clinical, and genetic risk factors (24). For instance, they considered BMI (27, 40, 41), behavioral factors such as breastfeeding (15, 19, 42) and hormone replacement therapy (HRT) use (28), biochemical data such as nipple aspirate fluid (NAF) cytology (43), radiological data such as BI-RADS features (26, 44) and mammographic density (27, 45, 46), and genetic variants such as single nucleotide polymorphisms (SNPs) (47, 48). Ericsson et al. (28) developed a prediction model and included HRT use, family history of breast cancer, menopausal status, and mammographic features into consideration. Based on data from a large cohort study, the AUC of their new model reached 0.71 compared to the AUC of 0.56 for the Gail model. Wu et al. (26) could significantly improve the discrimination by adding Genetic variants and mammographic features to the Gail model (from AUC of 0.59 to 0.73). Adding mammographic PD, BMI, and 18 SNPs to the Gail model improves the AUC from 0.55 to 0.62 (27). Attempts to improve risk assessment tools for breast cancer risk is ongoing, and scientists hope to reach an optimal tool for defining high-risk group and prevent breast cancer among this group.

Conclusion

The AUC of the Gail model was almost similar to that of other studies conducted in the USA and other countries. The Gail model has a modest discriminatory power to predict the risk of developing breast cancer among the Iranian female population. Cohort studies are still required to evaluate the validity of the model for the Iranian population.

Ethical considerations

Ethical issues (including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, redundancy, etc.) have been completely observed by the authors.

Acknowledgments

This study was a part of the M.Sc. thesis of the first author supported by Tehran University of Medical Sciences. The authors wish to thank Dr. Mitchell H. Gail, Senior Investigator of the Division of Cancer Epidemiology & Genetics, Biostatistics Branch at NCI, for his review and comments on the manuscript.

Conflict of interest

The authors declare that there is no conflict of interest.

References


Available at: [http://ijph.tums.ac.ir](http://ijph.tums.ac.ir)