Abstract

Background: Mammography is a valuable tool for early diagnosis of breast cancer in asymptomatic women. Considering the high prevalence of breast cancer in Iranian women and the low participation in mammography screening program, the purpose of this study was to investigate the factors affecting frequency of mammography screening in women over 40 years of age using zero-inflated Poisson regression models.

Methods: In this study, the required information about number of performing mammography in women’s lifetime, demographic characteristics and behavioral risk factors were obtained through interview based on a researcher-made questionnaire. To investigate the factors affecting mammography, zero-inflation Poisson regression models were performed using Win Bugs software.

Results: The mean (SD) age of women participating in this study was 49.87 (6.76). 77% of the participants have never undergone mammography, 8.9% once, 6.9% twice, 6.7% three times, and 0.5% more than three times. Age had a positive effect on the number of mammograms in the women who have performed mammograms at least once. Having a family history of cancer and breast cancer, middle compared to low economic status, higher compared to low education and menopause were significantly associated with lower probability of never performing mammography.

Conclusion: Given the relatively low participation of women in mammography, more facilities are needed for high risk women (aged 40-70) to facilitate the diagnosis process.

Keywords: Mammography, Breast cancer, Zero-Inflated, Poisson Regression, Bayesian Approach

Introduction

Breast cancer is the most common cancer among women, accounting for 23% of all cancers and 14% of mortalities (1). According to the World Health Organization, more than half a million mortalities occurred due to breast cancer in 2015 (2). In Iran, the results showed that the mortality rate from breast cancer from 1.97 per 100,000 in 2006 increased to 2.45 in 2010 (3). Unfortunately, it develops more than a decade earlier in Iranian women (4, 5) and about 70% of cases refer at advanced stages of the disease, which reduces the effectiveness of treatment and survival rate (6) so that the 10-year survival rate be less than 50% (7), while breast cancer is considered one of the treatable cancers, if diagnosed early (8). Although frequency of performing mammography...
mammography has increased dramatically in the last 30 years, it still remains below the recommended level in many societies (9). Although performing a screening mammography program is the best way to diagnose breast cancer at an early stage (10), studies show that only 6% of eligible women perform mammography on a regular basis during a ten-year period (11). In a study in the United States on women aged 24-69, almost 77% with private insurance completed mammograms (12). While in Iran, this rate was between 3% and 26% in various studies (13).

According to different cultures, there are various barriers to mammography, for example, in a review study on minority women in the United States, the most important barriers included low educational level and lack of health insurance, inadequate information on breast cancer screening programs, lack of physician’s advice, lack of trust in hospital and doctors, language barriers, and lack of transportation facilities (14). In a study conducted in Iran, factors such as fear of diagnosis, unawareness of screening programs, no free time, no breast problem, forgetfulness and negligence, and considering screening unnecessary were mentioned as the main factors for not performing mammography (13).

Generally, conventional regression modeling such as Poisson regression are used to identify factors associated with mammography screening in women. In the Poisson regression model, the goal is to model the average number of occurrences of the desired event based on independent variables. The assumption of the equality of mean of distribution with variance is the most important assumptions in the application of this model. This assumption limits dispersion of the obtained data; as a result, this model becomes more inefficient to fit the dispersed data and cause some problem such as inaccurate estimates of standard errors, and misleading p values. In fact, some count data contain a large number of zeros, which causes the data to be inappropriately adapted to the commonly used count regression models (Poisson, Negative Binomial). In such cases, alternative models such as negative binomial distribution or generalized Poisson distribution has been introduced, in which an additional parameter is included for explanation of data dispersion (15). But in the situation where there is many individuals with a count of 0, the zero-inflated Poisson model can be an alternative solution. Zero-inflated models have become fairly popular in the research literature because of applying these models need more advance statistical package. This model has been applied in a number of health related issues such as modeling DMFT index, number of falls in elderly, number of return to blood donation physical activity, and length of hospital stay (16-20). According to the review literature by the author this model has never been applied on mammography screening data.

Therefore, considering a significant percentage of women never undergo mammography in their lifetime, the main aim of this study was to examine the factors affecting the frequency of mammography with appropriate statistical models.

Methods
This was a cross-sectional study on women more than 40 years old referred to health care centers of Karaj, Iran. For calculating sample size required for this study, we used single population proportion formula With 95% confidence interval, 44 percent proportion (P) of women do mammography and 5% error (d), the sample size required was 380. For gathering data, we choose 10 from 20 health center and at each center questionnaire were filled from women more than 40 years.

Data collection form included questions about the number of mammography in women’s life-time, demographic characteristics and behavioral risk factors including drinking and smoking status.

In the Poisson regression model, the $i^{th}$ observation of the dependent (response) variable, the frequency of occurrence of an event, is a count variable that includes non-negative values. In Zero-inflated model, the use of zero inflated regression models is an alternative. In these models, it is assumed that $p_i$, the percentage of individual for whom the event does not occur, follow the Bernoulli distributions, and instead of the fitting of the commonly regression models, we use the mixture models, called zero inflated regression models (21).

The probability density function in the zero inflated models will be as follows:

$$P(Y_i|X_i) = \begin{cases} p_1 + (1-p_1) \Pr(Y_i = 0) & \text{if } y_i = 0 \\ (1-p_1) \Pr(Y_i = y_i) & \text{if } y_i = 1, 2, \ldots \end{cases}$$

Where $p_1$ is the probability of zero observation. If the $\Pr(Y_i = y_i)$ distribution is Poisson, we will have a Zero-Inflated Poisson regression, and if the distribution is considered negative binomial or generalized Poisson distribution, then there will be a Zero-Inflated Negative binomial and a Generalized Zero-inflated Poisson regression model. If $p_1 = 0$, the zero inflated models will become their common count regression models (22).

For data analysis in the Bayesian method we have to select a priori distribution for the model parameters and by integrating the data from these prior distributions and the probability function, the posterior distributions will be formed and the estimation of the model parameters will be based on these posterior distributions (23). Due to the complexity and large dimensions of the above-mentioned distributions, it is not possible to compute the posterior distribution of model parameters by analytical method. Therefore, Markov chain Monte Carlo methods are used to inference about model parameters. For this purpose, by sequential sampling of the complete conditional distributions of the parameters, we construct Markov chains using the Metropolis Hastings algorithm, where the distribution of these chains is an appropriate approximation of the posterior distribution of the model parameters (24).

For choosing the best covariate in the Zero-Inflated Model we used univariate analyze. In univariate analysis, the relationship between variables with frequency of mammograms was assessed by Chi-square, Fisher Exact test or one-way ANOVA. Normality assumption of continuous variables was assessed by Shapiro-Wilk test.

To fit the model, the program is written in Win Bugs software (25). After production of samples for the diagnosis of convergence, the Gelman-Rubin statistic was used to determine the appropriate burn-in period (26). Since the value of this statistic was less than 1.14 for all parameters,
Kavehie et al.

10,000 samples were suitable for checking convergence. As a result, 40,000 subsequent samples were considered as samples resulting from the posterior distribution of the parameters, samples were selected ten to one to reduce autocorrelation.

The smoothed histogram of the simulations indicates proper convergence of the Monte Carlo Markov chain algorithms. It is also seen that density function diagrams of the regression effects are almost symmetric, but the graphs of the density function λ are skewed. The stability of the graphs of the parameters’ effect on the specified domain and the absence of long-time fluctuations in these graphs show convergence of Monte Carlo Markov chain algorithms.

**Results**

The mean age of women were 49.87 (standard deviation (SD) = 6.76), with a minimum of 40 and maximum of 90 year old. The majority (99.5%) were married, 56.3% were under diploma. About 40% reported to be oral contraceptive users and 17.3% reported to be either smokers or alcohol consumers. In this study the outcome variable was the number of mammography sessions these women have undergone. Seventy seven percent of the participants had never performed mammography, 8.9% had performed it once, 6.9% twice, 6.7% three times, and 0.5% more than three times. Twenty seven percent of the participants reported cancer and 15.1% breast cancer in one of their family members or relatives. The result of univariate analysis is shown in table 1. The univariate analysis showed that age, education, menopause, cancer or breast cancer in relatives, an income were significantly associated with number of mammography screening. However, smoking, alcohol consumption and oral contraceptive usage were not significantly associated with the frequency of doing mammography.

The mean number of mammography was 0.45 with standard deviation of 0.92. Given the high percentage of zero in this data, it seemed that the use of zero inflated regression models was necessary for best fitting the data.

The results of the posterior distribution of parameters of the Poisson and Zero-Inflated Poisson models is shown in table 2. The 95% credible interval were based on 20,000 samples, a 10000-burn in period.

It is worth noting that a model with a smaller Deviance Information Criterion (DIC) has a better fit to data. By comparing these two models, it can be seen that the zero inflated Poisson model with a fitting criterion of 600.9 was better than the Poisson model with a DIC of 695.6. As a result, in the following, we only assessed the result of fitting the zero inflated Poisson model.

As shown in table 2, the parameter estimate for age (λ_i) is the number of doing mammography among women who had undertaken the screening at least once. Age was the only positive contributor, meaning that with increasing age, the number of mammograms increased. In this model, the effect of other variables has been investigated on parameters (p) that was the percent of women never has done mammography.

In other words, the presence of cancer or breast cancer in the family members or relatives, high income, college degree and menopause had a positive impact on performing mammography.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Zero</th>
<th>one</th>
<th>Two</th>
<th>Three</th>
<th>Forth</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, Mean (SD)</td>
<td>49.5 (6.6)</td>
<td>48.5 (6.6)</td>
<td>50.8 (4.9)</td>
<td>54.9 (7.9)</td>
<td>56 (11)</td>
<td>0.001</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Under educated</td>
<td>271 (87.1)</td>
<td>22 (61.1)</td>
<td>22 (78.6)</td>
<td>15 (55.6)</td>
<td>2 (100)</td>
<td>0.001</td>
</tr>
<tr>
<td>College degree</td>
<td>40 (12.9)</td>
<td>14 (38.9)</td>
<td>6 (21.4)</td>
<td>12 (44.4)</td>
<td>0 (0)</td>
<td>0.004</td>
</tr>
<tr>
<td>Menopause</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>180 (57.9)</td>
<td>19 (52.8)</td>
<td>8 (28.6)</td>
<td>8 (29.6)</td>
<td>1 (50)</td>
<td>0.001</td>
</tr>
<tr>
<td>Yes</td>
<td>131 (42.1)</td>
<td>17 (47.2)</td>
<td>20 (71.4)</td>
<td>19 (70.4)</td>
<td>1 (50)</td>
<td>0.001</td>
</tr>
<tr>
<td>Family history of cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>244 (78.5)</td>
<td>12 (33.3)</td>
<td>12 (42.9)</td>
<td>17 (63)</td>
<td>1 (50)</td>
<td>0.001</td>
</tr>
<tr>
<td>Yes</td>
<td>67 (21.5)</td>
<td>24 (66.7)</td>
<td>16 (57.1)</td>
<td>10 (37)</td>
<td>1 (50)</td>
<td>0.001</td>
</tr>
<tr>
<td>Family history of breast Cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>278 (89.4)</td>
<td>28 (77.8)</td>
<td>9 (32.1)</td>
<td>10 (37)</td>
<td>1 (50)</td>
<td>0.001</td>
</tr>
<tr>
<td>Yes</td>
<td>33 (10.6)</td>
<td>8 (22.2)</td>
<td>19 (67.9)</td>
<td>17 (63)</td>
<td>1 (50)</td>
<td>0.001</td>
</tr>
<tr>
<td>Drinking</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>296 (95.2)</td>
<td>35 (89.3)</td>
<td>25 (89.3)</td>
<td>27 (100)</td>
<td>2 (100)</td>
<td>0.41</td>
</tr>
<tr>
<td>Yes</td>
<td>15 (4.8)</td>
<td>1 (2.8)</td>
<td>3 (10.7)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0.36</td>
</tr>
<tr>
<td>Smoking</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>257 (82.6)</td>
<td>34 (94.4)</td>
<td>24 (85.7)</td>
<td>24 (88.9)</td>
<td>2 (100)</td>
<td>0.001</td>
</tr>
<tr>
<td>Yes</td>
<td>54 (17.4)</td>
<td>2 (5.6)</td>
<td>4 (14.3)</td>
<td>3 (11.1)</td>
<td>0 (0)</td>
<td>0.78</td>
</tr>
<tr>
<td>Oral contraceptive use</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>202 (65)</td>
<td>22 (61.1)</td>
<td>17 (60.7)</td>
<td>16 (59.3)</td>
<td>2 (100)</td>
<td>0.001</td>
</tr>
<tr>
<td>Yes</td>
<td>109 (35)</td>
<td>14 (38.9)</td>
<td>11 (39.3)</td>
<td>11 (40.7)</td>
<td>2 (100)</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Table 1. Frequency of Mammograms According to Personal and Social Characteristics Pf Women (n = 404)

Value are frequency (percent) unless otherwise indicated
The result shows that economic participation rate is significantly associated with poor participation in breast cancer screening (32%). One of the reasons for this difference is the age of the studied participants. Given that mammography is usually recommended for women over 40 and at younger ages it is recommended in case of symptoms during clinical examinations. In this study age has positive effect on doing mammography (33). However, the age of the women was 41. The mean number of doing mammography; \( p_i \), the probability of never performing mammography screening.

### Table 2. Parameter Estimation of the Poisson and Zero-Inflated Poisson Regression Models

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameter</th>
<th>Covariate</th>
<th>Estimate</th>
<th>SE</th>
<th>2.5 Percentile</th>
<th>97.5 Percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poisson</td>
<td>( \lambda_i )</td>
<td>Constant</td>
<td>-3.24</td>
<td>0.56</td>
<td>-4.32</td>
<td>-2.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Age</td>
<td>0.025</td>
<td>0.012</td>
<td>0.003</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cancer</td>
<td>0.63</td>
<td>0.17</td>
<td>0.28</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Breast cancer</td>
<td>0.56</td>
<td>0.18</td>
<td>0.21</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td></td>
<td>menopause</td>
<td>0.49</td>
<td>0.19</td>
<td>0.14</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Education</td>
<td>0.57</td>
<td>0.18</td>
<td>0.22</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td></td>
<td>income</td>
<td>0.63</td>
<td>0.17</td>
<td>0.3</td>
<td>0.98</td>
</tr>
<tr>
<td>Zero-Inflated Poisson</td>
<td>( \lambda_i )</td>
<td>Constant</td>
<td>-1.8</td>
<td>0.56</td>
<td>2.94</td>
<td>-0.84</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Age</td>
<td>0.04</td>
<td>0.01</td>
<td>0.022</td>
<td>0.061</td>
</tr>
<tr>
<td></td>
<td></td>
<td>History Cancer</td>
<td>0.04</td>
<td>0.01</td>
<td>0.022</td>
<td>0.061</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Family history of Breast cancer</td>
<td>-1.89</td>
<td>2.1</td>
<td>-3.76</td>
<td>-0.42</td>
</tr>
<tr>
<td></td>
<td></td>
<td>menopause</td>
<td>-0.81</td>
<td>0.4</td>
<td>-1.61</td>
<td>-0.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>education</td>
<td>-2.1</td>
<td>1.18</td>
<td>-3.68</td>
<td>-0.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td>income</td>
<td>-1.66</td>
<td>0.43</td>
<td>-2.58</td>
<td>-0.87</td>
</tr>
</tbody>
</table>

\( \lambda_i \), the mean number of doing mammography; \( p_i \), the probability of never performing mammography.

### Discussion

Considering that mammography is a valuable tool in early diagnosis of breast cancer, identifying the contributing factors for doing mammography among women is the goal of this study. In this study, 77% of the participants had never performed mammography. The result shows that this screening method does not have a high acceptance among Iranian women. Other Iranian studies reported a lower participation rate based on the age range of the participants, which varies from 39.8% to 1.3% in various studies (27-30). One of the reasons for this difference is the age of the studied participants. Given that mammography is usually recommended for women over 40 and at younger ages, it is recommended in case of symptoms during clinical examinations.

In this study age has positive effect on doing mammography in women who go for a mammography at least once. Whereas according to the studies in other countries age had negative effect; in a study conducted in Turkey, elder women had 1.1 fold greater risk of nonparticipation in mammography screening (31). In Korean women, age had significantly contributed to the poor participation in breast cancer screening programs (32). One of the reason for this difference is related to the modeling strategy for analyzing data. The other reason might be that the breast cancer diagnosis in Iranian women is one decade earlier than other countries.

Moreover, social and economic factors have significant effects, as women’s participation is very low in the underdeveloped countries; for example, in Jamaica, only 11.4% of women performed mammography at least once in the past five years (33). In Nigeria, only 1.6% had experience of mammography (34). In Malaysian female teachers older than 40, 13.6% performed at least one mammogram (35). In South Africa, 94.7% of the participants never used any breast cancer screening methods (36).

In underdeveloped countries, the participation rate is almost the same as Iran: 75% of women over 40 in Turkey (37), 76.9% in Taiwan (38), 56% in Singapore (39), and 69.1% in Qatar (40) had no experience of mammography, and in Brazil 55.6% have not performed mammography in the past two years (41). On the other hand, the participation rate is higher in the developed countries: 85% of women aged 50-70 in Greece, 59.9% of Japanese women aged 40-70 (2), and 65.9% of African women aged 50-74 residing in Australia have performed mammography every two years and only 19% of them did not undergo mammograms (42). In the United States, immigrant women’s participation in mammography varies based on different ethnic origins between 48.5% and 74.5% (43).

The finding of current study showed that economic condition has a positive impact on performing mammography. This finding in accordance with previous study (31). In a study in Switzerland, divided into two French and German regions, with free of charge and charged screening program, the results showed that in women aged 50-69, 77.8% and 34.9% performed mammography in the past two years, in the mentioned areas, respectively. This study revealed that free of charge mammography in a prosperous country like Switzerland causes two-fold increase in the mammography rate (44).

The results of this study showed that women’s participation in mammograms, in addition to cultural and social issues, is related to two important factors including age and income. Older age women tended to perform mammography higher than younger age groups (45), the most important reason for this finding can be the higher risk perceived by the older women. Also, higher participation rates are observed in higher-income countries and women with higher socioeconomic status in low and middle income countries (46).

Another important factor affecting women’s tendency to perform mammography was positive family history of cancer or breast cancer; these women are more likely to do mammogram because they feel more at risk. This result is also confirmed by other studies. In fact, women who found themselves at lower risk for developing cancer, would be reluctant to perform mammography (47, 48). In contrast, in Saudi Arabia where twenty-six percent of women reported a family history of breast cancer, women who had a family history of breast cancer reported lower mammography use (32%) than those without a family history of the disease (49).

In our study smoking did not have significant effect on mammography but in a study conducted in Korea, smoking is significantly associated with poor participation in breast cancer screening programs (31).
In our study menopausal has significant negative effect on performing mammography and in study done in Turkey, postmenopausal had 1.5 times the risk of women did not participate in the mammography screening in comparison premenopausal.

In our study, college degree women had more tendency to do mammography in comparison to high school diploma and undergraduate women. The similar result were reported in other studies (45-51).

Conclusion
This study revealed that a high percentage of women were never performed mammography in their lifetime. Considering the fact that mammography is the only acceptable tool for early diagnosis of breast cancer in women, the appropriate facilities should be provided for high-risk individuals (women aged 40-70) to increase the acceptability of this method.

Acknowledgements
None.

Ethical consideration
The protocol of this study has been approved by Institutional Review Board of Alborz University of Medical Sciences, Karaj, Iran.

Conflicts of interests
Authors declared no conflict of interest.

Funding
None.

References
28. Heidari Z, Mahmoudezadeh-Sagheb HR, Sakhavar N. Breast cancer