BREAST CANCER PREVENTION: IS IT POSSIBLE TO IMPROVE THE SELECTION BY GAIL MODEL USING THE FUZZY LOGIC METHODOLOGY? A RETROSPECTIVE STUDY

Sirlei Siani MORAIS[1]
Giuliano Mendes DUARTE[2]
Renato TORRESAN[2]
Cesar CABELLO[2]

- *ABSTRACT: Patients' reaction to chemotherapy is directly related to the state of their hormonal receptors. In most cases, studies on the chemoprevention of breast cancer use women with a not necessarily hormone-sensitive high risk of breast cancer (using the Gail Model). This study compared the capabilities of a new hybrid model and the Gail model in selecting women to chemoprevention. The objective was to determine the better method to select women who should be benefited with breast cancer chemoprevention. The new hybrid model was constructed using fuzzy sets classification of risk factors and crispy rules (constructed by translating physicians' perceptions of hormone-sensitive breast cancer's risk). This model considered age, age at menarche, number of previous biopsies, number of relatives affected by breast cancer, and age at first live birth. Since the new model had been developed, we calculated and compared the risks by Gail Model and by fuzzy hybrid model. The data used refer retrospectively to five years before the disease's diagnosis. The fuzzy hybrid model presented better results and improved significantly the accuracy of predicting hormone-sensitive breast cancer (20.5%). However, we believe that it is necessary to validate it with a large and prospectively sample before be used clinically.*

- *KEYWORDS: Breast cancer; hormone receptor; fuzzy logic; chemoprevention.*

1 Introduction

Tamoxifen, a risk-reduction medication for breast cancer in high-risk women, was approved in 1999 and has since been widely studied. Clinical trials have shown that tamoxifen treatment reduces the rates of tumors with positive estrogen receptor (ER+) by 30%-50% and has no effect on the incidence of estrogen receptor negative cancer (Fisher *et al.*, 1998; Cuzick *et al.*, 2003; Veronesi *et al.*, 2003; Fisher *et al.*, 2005; Vogel *et al.*, 2006; Cuzick *et al.*, 2007; Veronesi *et al.*, 2007). Other endocrine agents (raloxifene and aromatase inhibitors) used in high-risk women also reduce the incidence of ER+ but not

1 University of Campinas – Unicamp, Division of Statistics, Department of Obstetrics and Gynecology, CEP: 13083-887, Campinas, Brazil. E-mail: *sirleism@unicamp.br*

2 University of Campinas – Unicamp, Division of Oncology and Senology, Department of Obstetrics and Gynecology School of Medicine, CEP: 13083-887, Campinas, Brazil. E-mail: *giuliano.duarte@ig.com.br* / *torresan@terra.com.br* / *cabello@correionet.com.br*

ER- cancers (Cauley *et al.*, 2001; Goss and Strasser-Weippl, 2004). Tumors that are positive for an ER or progesterone receptor (PR) are considered to be hormone-sensitive, since the presence of a PR indicates a functional ER (Osborne, 1998).

The most commonly used model to assess breast cancer risk for the purpose of chemoprevention trials is the Gail model (Gail *et al.*, 1989; Fisher *et al.*, 1998; Costantino *et al.*, 1999; Cuzick *et al.*, 2003; Fisher *et al.*, 2005; Vogel *et al.*, 2006). Essentially, the Gail model considers simple markers: age, age at menarche, number of biopsies, presence of atypical hyperplasia (or lobular neoplasia), age at first live birth, number of close relatives with breast cancer, and race (Gail *et al.*, 1989; Costantino *et al.*, 1999). Most of these factors are associated with variances in the risk of hormone-sensitive breast cancer (and other markers) (Key and Pike, 1988; Cotterchio *et al.*, 2003; Althuis *et al.*, 2004; Ma *et al.*, 2006; Hines *et al.*, 2008).

Several factors influence the expression of receptors, such as the number of children (each birth reduces the risk of ER+ cancer by 11%) (Ma *et al.*, 2006); advanced age at the first live birth (increases the risk of ER+PR+ cancer by 27%); early menarche; breastfeeding (Althuis *et al.*, 2004; Ma *et al.*, 2006); body mass index (BMI) (Althuis *et al.*, 2004); and a family history of breast cancer (in pre-menopausal women this increases risk for ER- cancer) (Cotterchio *et al.*, 2003), and sometimes race (Hines *et al.*, 2008). Although the Gail model provides both an accurate estimate of breast cancer risk, and the parameters for estimating risk related to hormone-sensitive breast cancer, it has certain limitations (Clamp *et al.*, 2002). This model does not subdivide risk according to receptors. It may also underestimate the risk in relation to a family history of breast cancer, since it does not consider second and third degree relationships – a factor that a physician may certainly consider as a risk (Clamp *et al.*, 2002; Gail *et al.*, 2007; Veronesi *et al.*, 2007).

Recently, a new mathematical method has proposed to translate human thought in a way that can be programmed into computers: fuzzy logic (Zadeh, 1965; Sivanandam *et al.*, 2007; Massad *et al.*, 2008; Vineis, 2008). This method has shown better results than traditional statistical methods for some applications (Nakashima *et al.*, 2005; Castanho *et al.*, 2007; Sahan *et al.*, 2007; Ghazavi and Liao, 2008). In traditional logic, each person is either present or not in a single category (she/he does or does not belong to the category). We can summarize the idea of fuzzy logic by saying that a fuzzy set offers flexibility to the concept of belonging. In classical set theory one is able to classify subjects into classes through a binary process, that is, according to whether characteristic x does or does not belongs to set A. In fuzzy logic, on the other hand, the idea that characteristic x belongs to set A can be represented by a mathematical function known as a membership function of fuzzy set A (Massad *et al.*, 2008). This function reflects the certainty that allows one to classify this number (x) as A characteristic (Massad *et al.*, 2008).

Studies on breast cancer that have included fuzzy logic have produced good results (Nakashima *et al.*, 2005; Ghazavi and Liao, 2008; Huang *et al.*, 2010; Subbhuraam *et al.*, 2010). Good concordance of physicians thought patterns was observed in a breast density diagnosed study in which a fuzzy model demonstrated close concordance with the diagnostic given by radiologists (99.1% to 100%) (Bayram and Acar, 2007).

In such a context, predictive models that can classify the chance of hormone-sensitive breast cancer may be more useful and more efficient in determining the need for preventive treatment than the current classification models (Clamp *et al.*, 2002; Veronesi *et al.*, 2007).

In the current study, we applied fuzzy logic to classify continuous variables already established as risk factors (Key and Pike, 1988; Cotterchio *et al.*, 2003; Althuis *et al.*, 2004; Ma *et al.*, 2006; Hines *et al.*, 2008) and used this classification in a traditional base of rules to build a hybrid model that associates fuzzy and others concepts in order to assess (five years before diagnosis) the risk of hormone-sensitive breast cancer. To evaluate the effectiveness of this model to selected women for breast cancer chemoprevention, we also estimated the breast cancer risk through the Gail model (classified in high risk) and compared both sets of predictions to results based on anatomical-pathological results. This procedure (as outlined above) allowed us to evaluate how predictive models can be used to select women who might benefit from breast cancer chemoprevention with tamoxifen (or new agents).

2 Material and methods

2.1 Subjects

453 charts were analyzed through sequential retrospective review. The patients had been treated at the breast clinic of the Women's Integral Healthcare Center of the University of Campinas (Brazil) between January 2006 and June 2008. Only female patients between the ages of 30 and 70, with a confirmed diagnosis of any type of breast cancer, were retrospectively included. The ethical committee for research of the School of Medical Sciences, University of Campinas, approved the study and granted its consent. We excluded 4 women whose charts revealed that they had had chemotherapy previously, 5 who had had bilateral tumors, and 6 who had not been diagnosed with breast cancer for the first time. Those with charts containing incomplete data were also excluded. 31 charts did not show results concerning ER and PR, while 66 others lacked other data. 341 female subjects remained after all the exclusions. The sample size was defined by considering the prevalence of 70% for ER+ (Cuzick *et al.*, 2003), and a sensibility of 65% (with 10% of sample error), based on values obtained from validation studies of the Gail model (Rockhill *et al.*, 2001; Ulusoy *et al.*, 2009). A level of significance of 5% and a sample error for prevalence of 5% were assumed. The minimal size of the sample was thus $n = 268$.

2.2 Methods

Data covering five years retrospectively was collected based on characteristics analyzed by the Gail model and hormonal factors related to the receptors. These included age (age at diagnosis – 5 years), age at menarche, race, number of live births, age at first live birth, body mass index (BMI), menopause status, family history of breast cancer, previous biopsy and presence of atypical hyperplasia, and the results of anatomic-pathological examinations of estrogen and progesterone receptors. The women were classified according to the expression of the receptors. We assigned risk based on the two following models:

Gail Model

The Gail model is the risk model most commonly used in major studies on breast cancer chemoprevention trials (Cuzick *et al.*, 2003). It was developed in 1989 on a primarily caucasian population that participated in the Breast Cancer Detection and

Demonstration Project (BCDDP) (Gail *et al.*, 1989). In 1999 it was reformulated when the BCDDP's breast cancer incidence rates were replaced with estimates drawn from the Surveillance, Epidemiology, and End Results programs, which took into account risk estimates for African-American and Asian-American women (Costantino *et al.*, 1999). This model included six breast cancer risk factors, including age, age at menarche (less than 12 years, 12, 13, 14 or more), age at first live birth (nulliparous, less than 20, 20 to 24, 25 to 29, 30 or more), number of first-degree relatives with breast cancer (0, 1, 2 or more), number of biopsies (0, 1, 2 or more), and presence of atypical hyperplasia on a biopsy (yes or no). The statistical methodology was based on a prospective study with Cox proportional risk, calculating the risk of developing breast cancer in a lifetime, and within five years (Gail *et al.*, 1989; Costantino *et al.*, 1999). The five-year risk used in chemoprevention trials classified women as low risk: ≤ 1.66, and high risk: those with risk >1.66 (Fisher *et al.*, 2005). In our sample we evaluated the women retrospectively, calculated their risk, and classified risks as low and high.

Hybrid model

The risk assessment of the hybrid model was constructed from a union of fuzzy classifications from interval variables and a traditional base of rules, based on the physician's idea. The model can be broken down into four steps: 1) selection of variables; 2) construction of fuzzy sets; 3) definition of the base of rules; and 4) risk classification.

Step 1) The selection of variables

In the construction of the hybrid model, all variables of the Gail model were evaluated and risk factors demonstrated through meta-analysis studies and the expertise of the specialists. The variables included in the Gail model were: age (age at diagnosis – 5 years), age at menarche, age at first live birth (or nulliparous), previous biopsy with or without presence of atypical hyperplasia, race, and first-degree family history of breast cancer. We also considered a second- and third-degree family history of breast cancer, a highly relevant factor not considered by the Gail model (Clamp *et al.*, 2002). The variables included in the meta-analysis studies were those that showed like risk factors or risk limits, preferably those linked to estrogen: number of live births, menopause status, degree of family members with breast cancer, and breastfeeeding. To approximate this model to a fuzzy model, we also considered other observations made by the specialists. A group of three specialists (physicians) evaluated the variables described, and agreed with the list, but two of them considered BMI a variable that needed to be studied, so we included it as well.

Step 2) Construction of fuzzy sets

We used triangular and trapezoidal functions in order to determine the fuzzy sets of the continuous variables, following the equations of fuzzy sets. The general equations of fuzzy sets were obtained by using traditional functions of triangular (1) and trapezoidal sets (2). The specialists must determine the points a, b, c and d (Equations 1 and 2), for this they can use reported values in literature and their knowledge (Lc and Rc, 2006). In order to defined the values of a, b, c and d of fuzzy sets to construct the fuzzy sets of the model, the three specialists (the physicians) and the statistician assessed the classes of risk for developing hormone-dependent breast cancer through meta-analysis studies (Cotterchio *et al.*, 2003; Ma *et al.*, 2006; Hines *et al.*, 2008) as well as from variables

measured by the Gail model (Costantino *et al.*, 1999), and used the expertise of the specialists to assign fuzzy sets to continuous variables: age, age at menarche, age at first live birth, and BMI. The statistician who constructed the fuzzy sets worked individually with each of the specialists, using the software MATLAB, version 6.5, to input values of *a*, *b*, *c* and *d*. Afterwards, the fuzzy sets of each specialist were evaluated by the other two, and final sets were established by the consensus of at least two of them. The fuzzy sets produced by each specialist were very similar or identical, so consensus was easily achieved. Because the BMI did not change the risk (this will be explained more fully in item III) we have not provided this number. Therefore, the values of *a*, *b*, *c* and *d* (Equations 1 and 2) we determined using knowledge of specialists, and references values of literature, resulting in three equations for each variables (Equations 3 to 11).

Figure 1 -

$$\mu(x) = (\textit{triangular function}) \begin{cases} 0 & \textit{if} \quad x < a \\ \dfrac{x-a}{b-a} & \textit{if} \quad a \le x < b \\ \dfrac{c-x}{c-b} & \textit{if} \quad b \le x < c \\ 0 & \textit{if} \quad x > c \end{cases} \tag{1}$$

Figure 2 -

$$\mu(x) = (\textit{trapezoidal function}) \begin{cases} 0 & \textit{if} \quad x < a \\ \dfrac{x-a}{b-a} & \textit{if} \quad a \le x < b \\ 1 & \textit{if} \quad b \le x < c \\ \dfrac{d-x}{d-c} & \textit{if} \quad c \le x < c \\ 0 & \textit{if} \quad x > d \end{cases} \tag{2}$$

All fuzzy numbers were constructed before the data were collected and without observing the sample. They were based on the expertise of the specialists and data in the literature and the resulting fuzzy sets are presented in Figures 1 (age), 2 (age at menarche) and 3 (age at first live birth) (Massad *et al.*, 2008).

For example, according to the sets presented in Figure 1, a woman of 39 years old can be classified as either very young (*VY*), young (*Y*), or elderly (*E*), and the degree of pertinence for the three classes is calculated according to: $\mu_{VY}(x_1) = 0.1$; $\mu_Y(x_1) = 0.75$; $\mu_E(x_1) = 0$. Her maximum grade of pertinence is therefore 0.75, which corresponds to the "young" class. Based on Figures 1 to 3, the equations of pertinence grades for each fuzzy can be described like Equations 3 to 11.

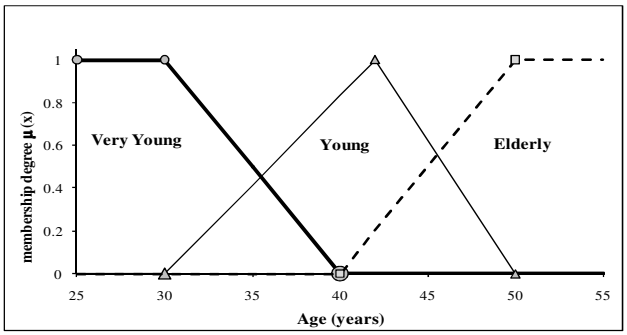

Figure 1 - *Fuzzy set numbers to statement age x = Age in years;*
$\mu_{\text{Very-Young}}(x)$; $\mu_{\text{Young}}(x)$; $\mu_{\text{Elderly}}(x)$.

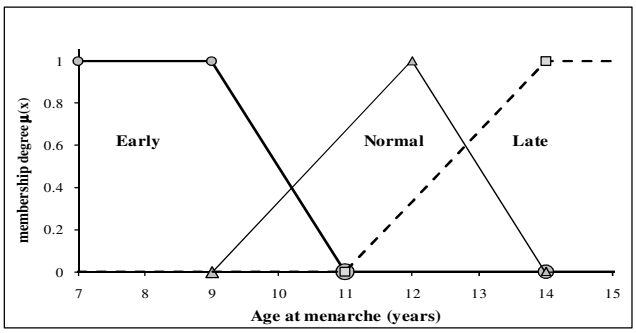

Figure 2 - Fuzzy set numbers to statement age at menarche: x = age at menarche in years;
$\mu_{\text{Early}}(x)$, $\mu_{\text{Normal}}(x)$, $\mu_{\text{Late}}(x)$.

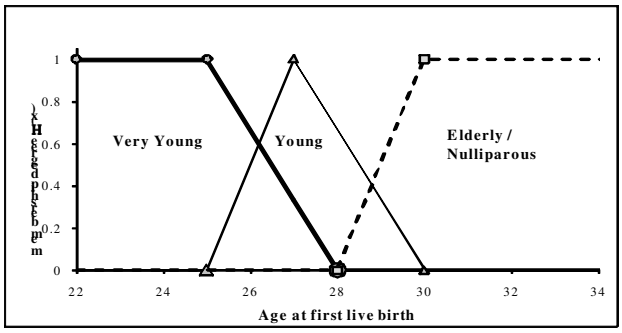

Figure 3 - Fuzzy set numbers to statement age at first live birth: x = age at first live birth in years
$\mu_{\text{Very-Young}}(x)$, $\mu_{\text{Young}}(x)$, $\mu_{\text{Elderly/Nulliparous}}(x)$.

Fuzzy set to age (years): x_1 = Age: [25,70] - Equations 3 to 5.

Very Young, Equation 2: a=b=25, c=30, d=40

$$\mu_{VY}(x_1) = \begin{cases} 1 & 25 < x_1 \leq 30 \\ \dfrac{40 - x_1}{10} & 30 < x_1 \leq 40 \\ 0 & x_1 > 40 \end{cases} \quad (3)$$

Young, Equation 1: a=30, b=42, c=50

$$\mu_{Y}(x_1) = \begin{cases} 0 & x_1 \leq 30 \\ \dfrac{x_1 - 30}{12} & 30 < x_1 \leq 42 \\ \dfrac{50 - x_1}{8} & 42 < x_1 \leq 50 \\ 0 & x_1 > 50 \end{cases} \quad (4)$$

Elderly, Equation 2: a=40, b=50, c=d=70

$$\mu_{E}(x_1) = \begin{cases} 0 & x_1 \leq 40 \\ \dfrac{x_1 - 40}{10} & 40 < x_1 \leq 50 \\ 1 & x_1 > 50 \end{cases} \quad (5)$$

Fuzzy set to age at menarche (years): x_2 = age at menarche: [7,20] Equations 6 to 8.

Early, Equation 2: a=b=7, c=9, d=11

$$\mu_{E}(x_2) = \begin{cases} 1 & 7 < x_2 \leq 9 \\ \dfrac{11 - x_2}{2} & 9 < x_2 \leq 11 \\ 0 & x_2 > 11 \end{cases} \quad (6)$$

Normal, Equation 1: a=9, b=12, c=14

$$\mu_{Y}(x_1) = \begin{cases} 0 & x_2 \leq 9 \\ \dfrac{x_2 - 9}{3} & 9 < x_2 \leq 12 \\ \dfrac{14 - x_2}{2} & 12 < x_2 \leq 14 \\ 0 & x_2 > 14 \end{cases} \quad (7)$$

Late, Equation 2: a=11, b=14, c=d=20

$$\mu_E(x_1) = \begin{cases} 0 & x_2 \leq 11 \\ \dfrac{x_2 - 11}{3} & 11 < x_2 \leq 14 \\ 1 & x_2 > 14 \end{cases} \quad (8)$$

Fuzzy set to age at first live birth: x_3 = age at first live birth: [10,70] Equations 9 to 11.

Very Young, Equation 2: a=b=10, c=25, d=28

$$\mu_E(x_3) = \begin{cases} 1 & 10 < x_3 \leq 25 \\ \dfrac{28 - x_3}{3} & 25 < x_3 \leq 28 \\ 0 & x_3 > 28 \end{cases} \quad (9)$$

Young, Equation 1: a=25, b=28, c=30

$$\mu_Y(x_3) = \begin{cases} 0 & x_3 \leq 25 \\ \dfrac{x_3 - 25}{3} & 25 < x_3 \leq 28 \\ \dfrac{30 - x_3}{2} & 28 < x_3 \leq 30 \\ 0 & x_3 > 30 \end{cases} \quad (10)$$

Elderly/Nulliparous, Equation 2: a=28, b=30, c=d=70

$$\mu_E(x_3) = \begin{cases} 0 & x_3 \leq 28 \\ \dfrac{x_3 - 28}{2} & 28 < x_3 \leq 30 \\ 1 & x_3 > 30 \end{cases} \quad (11)$$

Step 3) Base of rules definition

The idea of fuzzy models translates the thought process of physicians into mathematical rules (Massad *et al.*, 2008). Taking this into consideration, we constructed a base of rules that translates the physicians' perceptions by using a table with all combinations of variables: age (classified as: very young, young or elderly), age at menarche (early, normal and advanced), age at first live birth (very young, normal or elderly/nulliparous), BMI (normal and obese), number of previous biopsies (0,1 and ≥ 2), number of relatives affected by breast cancer (0, 1 and ≥ 2) and race (white or non-white). These combinations generated a table with 972 possibilities. Each of the physicians classified the combinations (rather than each woman) according to their risk – low, moderate and high – of developing hormone-dependent breast cancer. The variables –

breastfeeding, presence of hyperplasia, and menopause status – were also taken into account. The list contains a column to note whether these variables affected the risk in each combination. In most cases, the physicians were in concordance about the risk, but in cases where they were not, they reevaluated the categories together and reached a consensus. Many of the rules were redundant, as is often the case with traditional rules. Thus, for example, a woman whose age was classified as elderly, whose age at menarche was normal, who had one relative affected by breast cancer, who was nulliparous, and whose race was white, was considered at high risk of developing hormone-dependent breast cancer; in another rule, a woman was in the same classes of age (elderly), age at menarche (normal), relative affected by breast cancer (one), nulliparous, but whose race was non-white, was also considered at high risk. These two rules can be compressed in one rule if the race were not considered.

In order to simplify this base to a correct, minimal set of rules, we transformed the categories into binary numbers and simplified them through Boolean algebra, which we did with the help of the software BOOLE-DEUSTO (Boole-Deusto, 2003) and Karnaugh's map. Thus we reached the base of rules presented in Table 1.

Step 4) Risk classification

By this point we had established the linguistic classification of each variable as well as a classification system that included the other variables presented in Table 1. It is important to note that this model was constructed without access to data sets. The data set was applied only after the fuzzy hybrid model had been established in order to classify the risk of hormone-dependent breast cancer in each of the surveyed women (may be considered a validation data set) retroactively according to the following two steps:

1) Classification of the variables age, age at menarche, and age at first live birth, using the fuzzy sets in Figures 1 to 3;
2) Consideration of the linguistic classification obtained in step 1, the addition of the values of previous biopsies, the number of affected relatives, and the use of Table 1 to classify the risk.

Statistical Analysis

The actual response of hormonal receptors (according to the anatomical-pathological examination) was studied in two ways: ER+PR+ (both positives or at least one positive), and ER-PR- (both negatives), as previously explained. The women who had ER-PR- were considered control group – women that probably do not respond to chemoprevention (like the women that had a low risk to breast cancer). Whereas the fuzzy hybrid model was constructed according to the knowledge of the expert, without the collected data, the sample involved here could be considered a training data set. We compared the predicted responses of the Gail and the Fuzzy models by using statistics of validation, such as sensitivity, specificity, accuracy, positive predictive value, and negative predictive value. The results obtained were compared by means of the chi-squared test and Z test of proportion. Confidence intervals were calculated with a 5% level of significance. The softwares used were: SAS version 9.2, BOOLE-DEUSTO, and MATLAB 6.5 (Boole-Deusto, 2003)

Table 1 - Minimal base of rules, created by the physician and reduced* by Boolean algebra simplification.

Age	Age at Menarche	Previous Biopsies	No. of affected relatives**	Age at first live birth	Risk of hormone-sensitive breast cancer
Elderly	Normal	≥ 2	1	Elderly/Nulliparous	**High**
Elderly	Normal	anyone	0	anyone	**High**
Elderly	Early	0 or ≥ 2	0	anyone	**High**
Young	Late	0	0	anyone	**High**
Elderly	Late	1 or ≥ 2	0	anyone	**High**
Elderly	Late	0	0	Elderly/Nulliparous	**High**
Very Young	Early or Late	≥ 2	0	Elderly/Nulliparous	**High**
Young	Early	0	0	Young	**Low**
Young	Early	0	1	anyone	**Low**
Very Young	Late	1	anyone	anyone	**Low**
Very Young	Early	1	1	anyone	**Low**
Young	Late	≥ 2	1	Very Young or Young	**Low**
Very Young	Late	0	0	Very Young or Young	**Low**
Very Young	Normal	0	0 or 1	anyone	**Low**
Very Young	Late	0 or ≥ 2	1	anyone	**Low**
Very Young or Young	anyone	anyone	≥ 2	anyone	**Low**
All others cases the risk is moderate					

*Reduced was made with BOOLE-DEUSTO software / ** affected by breast cancer

3 Results

The average age for this study was 46.2 (9.8) years, and the average age at menarche was 13.0 (1.7) years. Most of the patients were Caucasian (87.7%); 13.2% were nulliparous; the average age at first live birth was 24 or less (65.9%); 44.9% were pre-menopausal, and most of the patients did not have any family members with breast cancer (75.3%). Among those who did have a relative with breast cancer, 75% reported a first-degree relationship. A previous biopsy had been carried out in 28.7% of the cases, and of these 43.9% had had atypical hyperplasia. The mean Gail risk was 2.43 (2.26), varying between 0.50 and 28.28. The hormonal receptors prevalence was 69.5% for ER+, 66.3% for PR+, 61.9% for both receptors (ER+andPR+) and 73.9% in at least one of the receptors (ER+and/orPR+) (Table 2).

In terms of positive prediction, the fuzzy hybrid model was more accurate than the Gail model for ER+PR+. Sensitivity was 89.3% (confidence interval = 85.5 to 93.1) in the fuzzy hybrid model and 49.6% (confidence interval = 43.4 to 55.8) in the Gail model ($p < 0.001$). In terms of positive predictive value, the highest values were observed in the fuzzy hybrid model, with values around 77%, despite the difference between these being

less than 5%. The negative predictive value of the fuzzy hybrid model was 43.7%, higher than the Gail model's 27.4%. With regard to this statistic, the fuzzy hybrid model was better than Gail's model, however, it was worse than the Gail model in terms of specificity: 53.9% (confidence interval = 43.6 to 64.3) for the Gail model as compared to 23.6% (confidence interval = 14.8 to 32.4) for the fuzzy hybrid model. This difference was a significant $p < 0.001$ (Table 3).

Table 2 - Distribution of social and hormonal characteristics of the women evaluated

Characteristics (n = 341)	n	%
Age, years		
< 30	15	4.4
30 to 49	202	59.2
≥ 50	124	36.4
Race		
White	299	87.7
Non-White	42	12.3
BMI, kg/m²		
< 25	132	38.7
25 to 30	113	33.1
> 30	96	28.2
Age at menarche, years		
9 to 14	279	81.8
> 14	62	18.2
No. of live births		
nulliparous	45	13.2
1	52	15.2
≥ 2	244	71.6
Age at first live birth, years		
≤ 24	195	65.9
25 to 29	59	19.9
≥ 30	42	14.2
Menopause Status		
peri-menopause	153	44.9
post-menopause	188	55.1
No. of affected relatives		
0	257	75.3
1	67	19.7
≥ 2	17	5.0
Grade of affected relatives		
first-degree	63	75.0
two or more degree	21	25.0
Previous Biopsy	98	28.7
Presence of atypical hyperplasia		
Yes	43	43.9
No	55	56.1
Positive estrogen receptor	237	69.5
Positive progesterone receptor	226	66.3

The general failure of models can be measured by 100% accuracy, and in terms of correct general classification, the fuzzy hybrid model presented better rate of prediction with an accuracy of 72.1% (confidence interval = 67.4 to 76.9), which was significantly different from that of the Gail model (50.7% confidence interval = 45.4 to 56.0, $p < 0.001$) (Table 3).

Table 3 - Evaluation of prediction by the Gail and fuzzy logic models, in relation to positivity of hormonal receptors (n = 341)

	ER+PR+ (n=252)	**ER-PR- (n=89)**	**Sensitivity (CI 95%)**	**Specificity (CI 95%)**	**Accuracy (CI 95%)**	**PPV**	**NPV**
Gail General Risk Assessment Model							
Risk > 1.66	125	41	49.6 (43.4-55.8)	53.9 (43.6-64.3)	50.7 (45.4-56.0)	75.3	27.4
Risk ≤ 1.66	127	48					
Fuzzy Model							
moderate	*135*	*41*	89.3 (85.5-93.1)	23.6 (14.8-32.4)	72.1 (67.4-76.9)	76.7	43.7
high	*90*	*27*					
moderate / high risk	225	68					
low risk	27	21					
Comparison of the performance of models:			p < 0.001	p < 0.001	p < 0.001*		

Chi-squared test / * Z test, PPV = Positive Predictive Value, NPV = Negative Predictive Value, CI = Confidence Interval

The subjects predicted to be at moderate risk by the hybrid model constituted the majority of the cases, and represented 51.6% of the 341 cases studied, 53.6% among ER+PR+ and 46.1% among ER-PR-. Of the 252 subjects that were ER+PR+, 43.2% were predicted as being at high/moderate risk in both models, 46.0% were predicted as being at high/moderate risk in the fuzzy hybrid model, but at low risk according to the Gail model, and 6.3% were predicted as being at high risk in the Gail model, while classified at low risk according to the fuzzy hybrid model. It may be concluded that fuzzy hybrid model failed in 6.3% of the cases where the Gail model was correct. However, the fuzzy hybrid model correctly predicted risk in 46.0% of the cases in which the Gail model did not and thus attained a 39.7% advantage. The same statistics were employed for 89 subjects with both negative receptors. In this case, a loss in predictive accuracy was observed with the fuzzy hybrid model. Nevertheless, in the 341 women evaluated, the fuzzy hybrid model presented 70 more women than the Gail model. These represented a profit in 20.5% of the sample (Table 4).

Table 4 - Evaluation of concordance in prediction by the Gail and fuzzy logic models, in relation to positivity of hormonal receptors (n = 341)

	Fuzzy Model			
	Moderate / High Risk		**Low Risk**	
	n	%	n	%
Gail Model				
ER+PR+ (n = 252)				
Risk > 1.66	109	43,2	***16***	***6,3***
Risk ≤ 1.66	***116***	***46,0***	11	4,5
ER-PR- (n = 89)				
Risk > 1.66	32	36,0	***9***	***10,1***
Risk ≤ 1.66	***39***	***40,4***	12	13,5
Profit by fuzzy model: 116 - 16 + 9 - 39 = 70 (20.5%)				

4. Discussion

The key to prevention strategies for breast cancer is selecting women likely to receive maximum benefit from the intervention with the least amount of risk (the drugs have side-effects). Alternative mathematical methods, such as the Gail model have been used to evaluate the usefulness of anti-hormonal substances such as tamoxifen in preventing breast cancer (Col *et al.*, 2002). Predictive models must therefore be evaluated in terms of both overall breast cancer risk, and specifically hormone-sensitive breast cancer risk (Veronesi *et al.*, 2007).

Gail's model is the most frequently used to predict positive hormone receptors in patients indirectly. Nonetheless it has limitations, such as low sensitivity (13.3%) and modest accuracy (58% and 53%) in certain validations. Furthermore, it takes into account only the underestimated risk of women with several affected relatives of second or third degree (Rockhill *et al.*, 2001; Clamp *et al.*, 2002; Ulusoy *et al.*, 2009). In our fuzzy hybrid model we have observed greater accuracy and a much higher degree of sensitivity. These results suggests that the use of this model, or an integration of this method, can lead to better results in breast cancer chemoprevention. Additional variables have been added to the fuzzy hybrid model but were not added to the Gail model using proportional hazards as the statistical approach: the reported differences in accuracy, sensibility and specificity can be smaller (or larger) than those reported, and should be evaluated by others studies.

The base of rules model, built on the knowledge of specialists, demonstrates consistence with factors also found in meta-analysis studies (Cotterchio *et al.*, 2003; Althuis *et al.*, 2004; Ma *et al.*, 2006; Hines *et al.*, 2008). According to the base of rules, women of an advanced age (fact showed by meta-analyses (Hines *et al.*, 2008)) or with

more than two relatives with breast cancer had a lower chance of hormone-sensitive breast cancer. This model also associated the number of previous biopsies with the greatest risk (Cotterchio *et al.*, 2003). Nulliparousness, advanced age at first live birth, and earlier menarche also increase risk (Cotterchio *et al.*, 2003; Althuis *et al.*, 2004; Ma *et al.*, 2006). Other factors, such as BMI, race, breastfeeding, and atypical hyperplasia, were not considered by minimal base of rules; published studies have not shown a convincing association (Cotterchio *et al.*, 2003; Althuis *et al.*, 2004; Ma *et al.*, 2006; Hines *et al.*, 2008) but future studies should examine these factors.

In terms of models used by physicians, we can say that both models present some flexibility. With the Gail model it is possible to change the default cut-off point used in chemoprevention trials (1.66%). The fuzzy hybrid model, on the other hand, provides easy flexibility since the moderate classification it offers can be understood either as constituting or not constituting a recommendation for chemoprevention. If a doctor chooses to be rigorous, he can recommend chemoprevention in moderate cases; if he chooses not to be so rigorous, he does not have to recommend chemoprevention in a case of moderate risk, or he may change the Gail cut-off point. Nevertheless the question "at what value?" will certainly crop up. The collateral effects of the drug used in chemoprevention as well as other variables of linguistic features, such as breast density classification (Bayram and Acar, 2007), can be weighted in fuzzy risk models. In traditional methods, however, this factor can be more complicated to implement.

In breast cancer chemoprevention, foreknowledge of a positive response and of collateral effects is crucial to the patients and physicians making the decision, and it is important to know their needs about the risk, however, there is few understanding of those with a family history of cancer (Wen and Gustafson, 2004). The patient's involvement in decision making about breast cancer prevention is important since it brings physical and psychological benefits (Col *et al.*, 2002; Gorla *et al.*, 2005). If a patient shows a low risk for receptors, other prophylactic treatments, such as surgery, must be considered and discussed between her and her physicians.

Although many chemoprevention studies must still evaluate this model, its positive results demonstrate that it offers another way to construct risk models. This model used in conjunction with the Gail model can help physicians to make decisions about chemoprevention. When considering chemoprevention treatment, physicians can use another model to evaluate breast cancer risk (such as the Gail or genetic models), and a model such as this one to evaluate the risk of hormone-dependent breast cancer. The question is: if a woman has a higher risk of developing breast cancer, what is the chance of her having a form of hormone-dependent breast cancer?

Predictive fuzzy logic models of the presence of hormonal receptors have not been described in published studies in cases of more specific comparisons, nor have other predictive models using different techniques (where sensitivity and/or specificity have not been determined (Gorla *et al.*, 2005)). Because it was observed that neither model that was studied had a very high degree of specificity, studies and methods for improve better predictions of breast cancer ER-PR- (negative response) are usually necessary.

Although the fuzzy hybrid model has generated accurate predictions, certain considerations about other methods should be taken into account. The Gail model is strict in determining the risk of breast cancer (Gail *et al.*, 1989; Costantino *et al.*, 1999) but is not specific with regard to hormonal receptors, and has not been validated in the case of Brazilian women.

The use of fuzzy logic in studies related to the breast demonstrate a level of accuracy above that obtained through our model (97%) (Nakashima *et al.*, 2005; Ghazavi and Liao, 2008), though various techniques employing fuzzy logic for the same set of breast cancers (associated with others techniques) have had an accuracy rate of between 80% and 97% (Nakashima *et al.*, 2005; Guo *et al.*, 2006; Bayram and Acar, 2007; Sahan *et al.*, 2007). This suggests that a different technique or fuzzy inference could improve the general performance of models.

As this is an initial retrospective study, some values can be seen as biased. The fuzzy hybrid model needs to be validated through a large sample, as do all models that deal with hormonal receptors.

Studies using more complete and sophisticated fuzzy methodologies, such as the total inference model of Mamdani or the Sugeno (Sivanandam *et al.*, 2007) method, can be carried out in order to improve the results presented here. Our model also considered menopausal status indirectly since most post-menopausal woman are of an advanced age, but this is a factor that does not apply in all cases. A model subdivided according to menopausal status, and one using a more thorough fuzzy method could serve as a continuation of this study.

Nowadays, genetics are recognized as a very important risk factor in breast cancer (and have attracted the attention of many people, including women). BRCA1 and BRCA2 deleterious mutations are strongly tied to breast cancer incidence in healthy women. This is a special population that needs to be targeted for prevention. However it is unclear whether women with a risk of inherent breast cancer are suitable for chemoprevention. Some retrospectives studies have shown that BRCA1 deletion mutation women have an elevated chance of tamoxifen chemoprevention failure. Further studies need to be performed in order to answer these questions, and the fusion of fuzzy and traditional methods may improve the result (King *et al.*, 2001; Dufloth *et al.*, 2005).

Since, generally speaking, the hydrid model is better than the Gail model in providing accurate predictions with regard to hormonal receptors, the Gail model should perhaps be adapted for hormonal receptors. It would thus be more helpful in identifying patients for whom treatment would be truly effective whenever chemoprevention is recommended (Gail *et al.*, 2007; Veronesi *et al.*, 2007). The increased attention paid to hormonal receptors is also likely to produce better results in the context of chemoprevention trials. Such predictions may be achieved by using fuzzy logic (Guo *et al.*, 2006; Phillips *et al.*, 2006; Ghazavi and Liao, 2008).

Conclusions

We conclude that using fuzzy classification and the expertise of a specialist to define risk may improve the selection of candidates for treatment in the categories studied, as demonstrated by the significant improvement in results achieved by the fuzzy hybrid model in comparison to the standard model. Although this is simply a preliminary study, and a validation in a large sample is recommended, results have shown an improvement in predictive accuracy with respect to other methods with relatively low specificity. Fuzzy logic may thus be more efficient than other models when it comes to avoiding cases of breast cancer in further prevention trials. The establishment of other methods relying on fuzzy logic (technically more sophisticated) in models testing

hormone-sensitive cancer may improve accuracy further. Fuzzy logic shows promise of positive results in clinical diagnosis. Risk is easily accessed by anyone; nevertheless this is a tool to aid physicians who are deciding whether or not to recommend chemoprevention to women.

MORAIS, S. S.; DUARTE, G. M.; TORRESAN, R.; CABELLO, C. Prevenção do Câncer de Mama: É possível melhorar a seleção pelo modelo de Gail usando a metodologia de lógica fuzzy? Um estudo retrospectivo. *Rev. Bras. Biom.*, São Paulo, v.29, n.3, p.416-434. 2011.

- RESUMO: A resposta de pacientes à quimioterapia está diretamente associada ao estado dos receptores hormonais. Na maioria dos casos, os estudos sobre quimioprevenção de câncer de mama selecionam as mulheres com um risco do mesmo, mas não necessariamente hormônio dependente (utilizando o modelo de Gail). Este estudo comparou as capacidades de um novo modelo híbrido e do modelo de Gail, em selecionar mulheres para quimioprevenção. O objetivo foi determinar o melhor método para selecionar mulheres que deveriam ser beneficiadas com a quimioprevenção do câncer de mama. O novo modelo híbrido foi construído utilizando a classificação dos fatores de risco em conjuntos fuzzy e regras crispy (transformando a percepção de clínicos sobre o risco de câncer de mama hormônio dependente). Este modelo considerou idade, idade à menarca, número de biópsias anteriores, número de parentes com câncer de mama, e idade ao nascimento do primeiro filho vivo. A partir do novo modelo desenvolvido, os riscos dos modelos de Gail e do novo modelo híbrido foram calculados e comparados. Os dados utilizados referem-se retrospectivamente a cinco anos antes do diagnóstico da doença. O modelo híbrido fuzzy apresentou melhores resultados e melhorou significativamente a acurácia na previsão de câncer de mama hormônio dependente (20,5%). No entanto, acreditamos que seja necessária a validação em uma amostra prospectiva e maior antes de ser usado clinicamente.
- PALAVRAS-CHAVE: Cancer de mama; receptor hormonal; lógica difusa; quimioprevenção.

References

ALTHUIS, M.; FERGENBAUM, J. H.; GARCIA-CLOSAS, M.; BRINTON, L. A.; MADIGAN, M. P.; SHERMAN, M. E. Etiology of hormone receptor-defined breast cancer: a systematic review of the literature. *Cancer Epidemiol. Biomark. Prev*., Philadelphia, v.13, n.10, p.1558-1568, 2004.

BARROS, L. C.; BASSANEZI, R. C. *Tópicos de lógica fuzzy e biomatemática*. Campinas: UNICAMP/IMECC, 2006. 344 p.

BAYRAM, B.; ACAR, U. An approach to the detection of lesions in mammograms using fuzzy image processing. *J. Int. Med. Res*., Worthing, v. 35, n. 6, p. 790-795, 2007.

BOOLE-DEUSTO. BOOLE-DEUSTO - Digital Logic Analysis and Design Environment. Desenvolvido por: BASILIO, B. S. Deusto-ES: Zubía JG; Martínez JS**:** International Conference on Microelectronic Systems Education. 2003. *Apresenta programa para simplificação de mapas de Karnaugh*. Disponível em: < http://paginaspersonales.deusto. es/zubia/ >. Acesso em: 30 nov. 2008.

CASTANHO, M.; BARROS, L. C.; YAMAKI, A.; VENDITE, L. L. *Fuzzy receiver operating characteristic curve*: an option to evaluate diagnostic tests. IEEE Trans. Inf. Technol. Biomed., Piscataway , v. 11, n. 3, p. 244-250, 2007.

CAULEY, J.; NORTON, L.; LIPPMAN, M. E.; ECKERT, S.; KRUEGERS, K. A.; PURDIE, D. W.; FARRERONS, J.; KARASK, A.; MELLSTROM, D.; NG, K. W.; STEPAN,J. J.; POWLES, T. J.; MORROW, M.; COSTA, A.; SILFENS, S. L.; WALLS, E. L.; SCHMITT, H.; MUCHMORE, D. B.; JORDAN, V. C. Continued breast cancer risk reduction in postmenopausal women treated with raloxifene: 4-year results from the MORE trial. Multiple outcomes of raloxifene evaluation. *Breast Cancer Res. Treat.*, The Hague, v.65, n.2, p.125-134, 2001.

CLAMP, A.; DANSON, S.; CLEMONS, M. Hormonal risk factors for breast cancer: identification, chemoprevention, and other intervention strategies. *Lancet Oncol*., London, v.3, n.10, p.611-619, 2002.

COL, N.; GOLDEBERG, G. J.; ORR, R. K.; ERBAN, J. K.; FORTIN, J. M. Survival impact of tamoxifen use for breast cancer risk reduction: projections from a patient-specific Markov model. *Med. Decis Mak*., Thousand Oaks, v. 22, n.5, p.386-393, 2002.

COSTANTINO, J.; GAIL, M. H.; PEE, D.; ANDERSON, S.; REDMOND, C. K. BENICHOU, J.; WIEAND, H. S. Validation studies for models projecting the risk of invasive and total breast cancer incidence. *J. Natl. Cancer. Inst.*, Bethesda, v.91, n.18, p.1541-1548, 1999.

COTTERCHIO, M.; KREIGER, N.; THEIS, B.; SLOAN, M.; BAHL, S. Hormonal factors and the risk of breast cancer according to estrogen- and progesterone-receptor subgroup. *Cancer Epidemiol. Biomark. Prev.***,** Philadelphia, v.12, n.10, p.1053-1060, 2003.

CUZICK, J.; CUZICK, J.; FORBES, J. F.; SESTAK, I.; CAWTHORN, S; HAMED, H.;HOLLI, K.; HOWELL, A. Overview of the main outcomes in breast-cancer prevention trials. *Lancet***,** London, v.361, n.9354, p.296-300, 2003.

_____. Long-term results of tamoxifen prophylaxis for breast cancer--96-month follow-up of the randomized IBIS-I trial. *J. Natl. Cancer Inst.***,** Bethesda, v.99, n 4, p.272-282. 2007.

DUFLOTH, R. M.; CARVALHO, S.; HEINRICH, J. K.; SHINZATO, J. Y.; SANTOS, C. C. dos; ZEFERINO, L. C.; SCHIMITT, F. Analysis of BRCA1 and BRCA2 mutations in Brazilian breast cancer patients with positive family history. *Sao Paulo Med. J.***,** São Paulo, v.123, n.4, p.192-197, 2005.

FISHER, B.; CONSTANTINO, J. P.; WICKERHAM, D. M.; CECCHINI, R. S.; CRONIN, W. M.; ROBIDOUX, A.; BEVERS, T. B.; KAVANAH, M. T.; ATKINS, J. N.; MARGOLESE, R. G.; RUNOWICZ, C. D.; JAMES, J. M.; FORD, L. G.; WOLMARK, N. Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. *J. Natl. Cancer Inst.***,** Bethesda, v.90, n.18, p.1371-1388, 1998.

_____. Tamoxifen for the prevention of breast cancer: current status of the National Surgical Adjuvant Breast and Bowel Project P-1 study. *J. Natl. Cancer Inst.***,** Bethesda, v.97, n.22, p.1652-1662, 2005.

GAIL, M.; ANDERSON, W. F.; GARCIA-CLOSA, M.; SHERMAN, M. E. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J. Natl.Cancer Inst.***,** Oxford, v.81, n.24, p.1879-1886, 1989.

_____. Absolute risk models for subtypes of breast cancer. *J. Natl. Cancer Inst.*, Oxford, v.99, n.22, p.1657-1659, 2007.

GHAZAVI, S.; LIAO, T. Medical data mining by fuzzy modeling with selected features. *Artif. Intell. Med.*, Amsterdam, v.43, n.3, p.195-206, 2008.

GORLA, S.; HOU, N.; ACHARYA, S.; RADEMAKER, A.; KHAN, S.; STARADUB, V.; MORROUW, M. A predictive model for the development of hormone-responsive breast cancer. *Ann. Surg. Oncol.*, New York, v.12, n.1, p.48-56, 2005.

GOSS, P.; STRASSER-WEIPPL, K. Prevention strategies with aromatase inhibitors. *Clin. Cancer Res.*, Philadelphia, v.10, n.1 pt.2, p.372S-379S, 2004.

HINES, L. M.; RISENDAL, B.; SLATTERY, M. L.; BAUGARTNER, K. B.; GIULIANO, A. R.; BYERS, T. Differences in estrogen receptor subtype according to family history of breast cancer among Hispanic, but not non-Hispanic White women. *Cancer Epidemiol. Biomark. Prev.*, Philadelphia, v.17, n.10, p.2700-2706, 2008.

HUANG, M. L.; HUNG, Y. H.; LEE, W. M. RK, L.; WANG, T. H. Usage of case-based reasoning, neural network and adaptive neuro-fuzzy inference system classification techniques in breast cancer dataset classification diagnosis. *J. Med. Syst.*, New York, 2010.

KEY, T.; PIKE, M. The role of oestrogens and progestagens in the epidemiology and prevention of breast cancer. *Eur. J. Cancer Clin. Oncol.*, New York, v.24, n.1, p.29-43, 1988.

KING, M., WIEAND, S.; HALE, K.; LEE, M.; WALSH, T.; OWENS, K.; TAIT, J.; FORD, L.; DUNN, B. K.; CONSTANTINO, J.; WICKERHAM, L.; WOLMARK, N.; FISHER, B. Tamoxifen and breast cancer incidence among women with inherited mutations in BRCA1 and BRCA2: National Surgical Adjuvant Breast and Bowel Project (NSABP-P1) breast cancer prevention trial. *JAMA*, Schaumburg, v.286, n.18, p.2251-2256, 2001.

MA, H.; BERNSTEIN, L.; PIKE, M. C.; URSIN, G. Reproductive factors and breast cancer risk according to joint estrogen and progesterone receptor status: a meta-analysis of epidemiological studies. *Breast Cancer Res.*, v.8, n.4, p.R43, 2006.

MASSAD, E.; ORTEGA, N.; STRUCHINER, C. Fuzzy Logic in Action: Applications in Epidemiology and Beyond. In: KACPRYZK, J. (Ed.). 1. *Basic Concepts of Fuzzy Sets Theory*. Berlim - German: Springer-Verlag, 2008. p.11-38.

NAKASHIMA, T.; SCHAEFER, G.; YOKOTA, Y.; ISHIBUCHI, H. Weighted fuzzy classification with integrated learning method for medical diagnosis. *Conf. Proc. IEEE. Eng. Med. Biol. Soc.*, v.6, p.5623-5626, 2005.

OSBORNE, C. Steroid hormone receptors in breast cancer management. *Breast Cancer Res. Treat.*, New York, v.51, n.3, p.227-238, 1998.

PHILLIPS, M.; CATANEO, R. N.; DITKOFF, B. A.; FISHER, P.; GREENBERG, J.; GUNAWARDENA, R.; KWON, C. S.; TIETJE, O.; WONG, C. Prediction of breast cancer using volatile biomarkers in the breath. *Breast Cancer Res. Treat.*, v.99, n.1, p.19-21, 2006.

ROCKHILL, B.; SPIEGELMAN, D.; BYRNE, C.; HUNTER, D. J.; COLDITZ. G. A.

Validation of the Gail et al. model of breast cancer risk prediction and implications for chemoprevention. *J. Natl. Cancer Inst.***,** Oxford, v.93, n.5, p.358-366, 2001.

SAHAN, S.; POLAT,K; KODAZ, H.; GUNES, S. A new hybrid method based on fuzzy-artificial immune system and k-nn algorithm for breast cancer diagnosis. *Comput. Biol. Med.***,** Oxford, v.37, n.3, p.415-23, 2007.

SIVANANDAM, S.; SUMATHI, S.; DEEPA, S. *Introduction to fuzzy logic using MATLAB*. New York-US: Springer, 2007. 409p.

SUBBHURAAM, V. S.; NG, E. Y. K.; KAW, G.; ACHARYA, R.; CHONG, B. K. Evaluation of the efficiency of biofield diagnostic system in breast cancer detection using clinical study results and classifiers. *J. Med. Syst.*, 2010.

ULUSOY, C.; KENEKCI, I.; KOSE, K.; AYDINTUG, S.; CAM, R. Applicability of the Gail model for breast cancer risk assessment in Turkish female population and evaluation of breastfeeding as a risk factor. *Breast Cancer Res. Treat.*, New York, v.120, n.2, 2009.

VERONESI, U.; MAISONNEUVE, P.; ROTMENSZ, N.; BONANNI,B.; BOYLE, P.; VIALE, G.; COSTA, A.; SACCHINI, V.; TRAVAGLINI, R.; DAIUTO, G.; OLIVIERO, P.; LOVISON, F.; GUCCIARDO, G.; TURCO, M. R. del; MURACA, M. G.; PIZZICHETTA, M. A.; CONFORTI, S.; DECENSI, A. Italian randomized trial among women with hysterectomy: tamoxifen and hormone-dependent breast cancer in high-risk women. *J. Natl. Cancer Inst.***,** Oxford, v.95, n.2, p.160-165, 2003.

_____. Tamoxifen for the prevention of breast cancer: late results of the Italian Randomized Tamoxifen Prevention Trial among women with hysterectomy. *J. Natl. Cancer Inst.***,** Oxford, v.99, n.9, p.727-737, 2007.

VINEIS, P. Methodological insights: fuzzy sets in medicine. *J. Epidemiol. Commun. Health***,** London, v.62, n.3, p.273-278, 2008.

VOGEL, V. G.; COSTANTINO, J. P.; WICKERHAM, D. L.; CRONIN, W. M.; CECCHINI, R. S.; ATKINS, J. N.; BEVERS, T. B.; FEHRENBACHER, L.; PAJON JR, E. R.; WADE III, J. L.; ROBIDOUX, A.; MARGOLESE, R. G.; JAMES, J.; LIPPMAN, S. M.; RUNOWICZ, C. D.; GANZ, P. A.; REIS, S. E.; McCASKILL-STEVENS, W.; FORD, L. G.JORDAN, V. C.; WOLMARK, N. Effects of tamoxifen vs raloxifene on the risk of developing invasive breast cancer and other disease outcomes: the NSABP Study of Tamoxifen and Raloxifene (STAR) P-2 trial. *JAMA***,** Schaumburg, v.295, n.23, p.2727-2741, 2006.

WANG & LU WANG, Y-J.; LU, S-X. Breast ultrasound image enhancement using fuzzy logic. *Ultrasound Med. Biol.***,** Philadelphia, v. 32, n. 2, p. 237-47 feb. 2006. Disponível em: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5207675 >.

WEN, K. Y.; GUSTAFSON, D. H. Needs assessment for cancer patients and their families. *Health Qual. Life Outcomes***,** London, v.2, p.11, 2004.

ZADEH, L. Fuzzy sets. *Inf. Control***,** Maryland Heights , v.8, p.15, 1965.

Received in 08.07.2011

Approved after revised in 20.10.2011