

# **Experience in appliance of the modern computer technologies in order to forecast the breast cancer's emergence within the conditions of the general medical network**

R.K. Tashchiyev<sup>1</sup>, V.T. Barateli<sup>2</sup>

<sup>1</sup>P.I. Shupik National medical academy of postgraduate education, Kyiv, Ukraine

<sup>2</sup>Novoselytsya Central District Hospital, Novoselytsya, Ukraine

**Summary.** The shortcomings in breast cancer's identification, as well as the ways to address its secondary prevention and its early detection are analyzed within this paper. A self-educational program is developed based on the neural network in order to calculate the risk of breast cancer's emergence and to create the risk's groups on cancer's occurrence.

**Key words.** Breast cancer, neural network, risk groups on emerging of the breast cancer, early detection of the breast cancer.

**Relevance.** Breast cancer (BC) is one of the most common women's tumor diseases. Moreover, the breast cancer incidence has been steadily increasing. During the last 30 years the global sickness rate increased from the 641000 cases in 1980 to 1643000 in 2010. The annual growth rate is equal to 3.1% [1]. According to the World Cancer Report (2008) there were 458503 cases of death from this disease worldwide in 2008. Among women older than 45 years, breast cancer is the most common cause of mortality, especially in the United States. While in Russian Federation the breast cancer disease is on the first place among all the female cancer diseases. Moreover, there is a slight growth in morbidity and mortality in among working-age population. The breast cancer's morbidity rate in Russia in 2011 was equal to 45.2 per 100 000 of the female population, with the mortality rate of 16.5 [2]. In Ukraine the breast cancer disease is on the first place, with an annual growth of 2.5-3%. In 2011 the morbidity rate was 19.6%, while the

mortality was equal to 20.2%. The breast cancer's incidence rate during the period of 2000-2010 increased from the 58.2 to 69.8 per 100 000 of the population. The mortality increased from 30.2 to 32.0 per 100 000 of the population during the same period of time, and is on the second place after lung, trachea and bronchi cancer. In the structure of incidence of malignant tumors among women under 30 years old, the breast cancer rate itself is equal to 7.2% [3].

It is also worth mentioning that the breast cancer disease is a multifactorial disease, which is characterized by a variety of clinical forms, as well as by the extreme prognostic diversity. At the same time the breast cancer's development, as of the hormone dependent tumor, is mostly based on different endocrine disorders arising at women with the irregularities in functioning of the ovarian, thyroid, adrenal, hypothalamic-pituitary systems'. [4, 5, 6]. Literature's data indicates that the first place among the precancerous breast diseases in recent years is taken by the fibrocystic disease (FCD) [7].

It is also highlighted [8] that the problem in diagnoses of the no palpable forms of the breast cancer is still unsolved and thus the non-invasive forms of breast cancer's detection are down to a random finding.

In recent years the methods of artificial neural network's construction is usually used for solving problems that require mathematical modeling within the problems of images recognition. Application of these methods proved to be effective in solving the problems of diagnosis, classification and prediction in medicine [9]. Based on the introduced to the program data the neural network calculated the risk's index for the respective object and issues the result in a form of the certain number in the range from 0 to 100. Thus, if the value is the closest to the upper limit, it can be asserted, that for this object the risk is considered to be very big.

In this respect it seems relevant to create and emerge into the clinical practice of the general medical network of the automated modern breast cancer risks' calculation system.

**Objective:** Creation of algorithm, which operates based on the neural network modeling, in order to calculate the index of individual risk of breast cancer, to divide examined women into risk groups for further clinical supervision.

**Research materials:** to study specific risk factors for breast cancer, we examined 360 women with clinical records of breast cancer, 270 women are registered with FCD, 90 women without breast pathology. After the risk factor's identification, which is typical for female population of the region, we examined 360 randomly picked clinically healthy women during the passing of the medical examination. In order to test the reliability of the obtained data from the usage the developed algorithm, we examined 360 women with clinical records of fibrocystic breast disease (FCD). All examined women were from 18 to 35 years old. The surveys were conducted based on the CI "Novoselytsya CDH" of Chernivtsi region.

**Methods of research:** clinical, laboratory and instrumental methods (mammography, ultrasonography). Method of questionnaire was used, with a purpose to identify criteria for the formation of "risk groups".

Methods of mathematical models' constructing were applied to establish links between the characteristic values obtained during the survey and the risk of breast cancer, namely: methods for constructing artificial neural network models, methods for constructing logistic regression models, as well as the program "oncologist" is based on a neural network.

**Results and discussion.** The majority of questioned women, which have clinical records in breast cancer – 297 (82.5%) had different risk factors or background diseases that could contribute to the development of breast cancer. The relation of risk factors' quantitative attributes and their dynamics within the different age groups was traced back in accordance with 297 profiles. A combination of up to 2 risk factors at the same examined woman with an age before 30 years have been noticed, some of them had up to 6 features in the group age of up to 40 years and 9 features in the group of age before 50 years, and stabilizing at 8 features within the

group of women older than 60 years. The number of risk factors from 1 to 4 were determined at 221 ( $74,4\pm 0,01\%$ ) women. A combination of 5 features at the same woman was marked at 27 ( $9,1\pm 0,01\%$ ), as well as six features at 19 ( $6,4\pm 0,01\%$ ). The biggest combination of risk factors was determined at 116 women within the age range from 41 to 50 years. The least of women (0.4%) indicated a pregnancy after 40 years, and most (17.7%) had thyroid disease (various forms of endemic goiter).

The typical features of risk factors were determined for each age group, which constituted the specific weight. At the same time the differences of these features within the various age groups are statistically significant ( $p < 0,001$  by criterion of  $\chi^2$ ). Among the 1016 features of risk factors, which were determined at 297 women, the first place was taken by variety forms of endemic goiter in 180 ( $17,7\pm 0,01\%$ ) of cases. Moreover, the same thing was observed within the age group between 21-30 years. Ischemic heart disease was determined in 161 cases ( $15,8\pm 0,04\%$ ) among all the age groups. Overweight issues were observed in 104 cases ( $10,2\pm 0,03\%$ ), liver diseases and biliary tracts in 88 cases ( $8,7\pm 0,1\%$ ), premenstrual mastodon in 60 cases (5.9%). Others risk factors from the one determined within the questionnaire were equal to 0.6% to 4%. The regularity of the monthly cycles during the life was determined only by  $43,9\pm 0,03\%$  of examined women. 43.9% of women stated the onset of the menopause after 50 years.  $20,3\pm 0,01\%$  of the female examinees gave birth to their first child after 26 years.  $46,0\pm 0,03\%$  of women stated the duration of breast feeding their last child was 4 months. During the conducted research we identified that breast cancer was more frequently found at the women, which had started their sex life at 17-18 years, namely  $68,9\pm 0,02\%$ . Also, the  $53,9\pm 0,02\%$  of women with breast cancer was diagnosed with FCD. In  $69,7\pm 0,08\%$  number of cases the patients with breast cancer were determined with an incensement of the body weight index (more than 30). The significance of input features of risk, after examination of the group of breast cancer patients, is collected in Table. All the questionnaire's data and

supplement methods of research were introduced into the developed program. During the second stage we examined 360 clinically healthy women. They were randomly chosen among the population who had been passing their oncology examination. Of the 360 polled women only 97 (26.9±0.06%) did not state at least one of the feature of the risk factors, in its turn it reflects the general unfavorable situation in the state of physical health of women in the studied region.

*Table*

*Significance of the input risk factors of breast cancer emergence*

Feature	Significance, %
Age	98
Premenstrual mastodon	89
Fibrocystic breast changes	88
Chronic inflammation of the appendages	73
Related thyroid diseases	69
Nipples discharges	64
Ovarian pathology	63
Menstrual cycle disorder	62
Tumors in other family members	62
Successive breast cancer	60
Uterus fibromyoma	57
Mastitis	57
Late menopause	55
Overweight	50
Lack of sex life	49
Diabetes	47
More than 5 abortions	42
Diminished libido	39
Ischemic heart disease	38
Liver and biliary tracts diseases	34
Giving birth to child with the weight of more than	32
Late onset of the menstrual cycle	31
Breast injury	30

Early onset of the menstrual cycle	28
Lack of pregnancy and labors	28
Late labors – after 28 years	21
Pregnancy after 40 years	19

This confirms the necessity to build a rational system of preventive examination for early detection of both benign and malignant diseases of the breast cancer in outpatient general medical network i.e. prospects of application of the developed by us medical technology selective screening (SSc). Women who received the index which was equal to 60 or more were included to the high risk group based on the conduction of analyzes of the every single questionnaire. The medium risk group of breast cancer emergence was constituted from the women who received the index of 30-60. The low risk group was made from women with the index of lower than 30.

The program used every single feature during the process of questionnaire analyzing, as well as the combination with the others features in a particular age group. Thus, for example, if the woman from the age group from 21-30 was not pregnant and didn't have labor, didn't have any sex life and didn't have any others features of the risk factors, than she was not included to the "group risk". All the above mentioned features were interpreted as the features for breast cancer's development. Women, who stated in their questionnaire only one factor, without a significant importance for the breast cancer's development, such as diminished libido, ischemic heart disease, overweight, diabetes, were not included to the "risk group" as well. In cases where these factors were in combination with others they contributed to the development of precancerous pathology of breast and women were included in the "risk group".

12 women were included to the "high risk group" among the 360 examined clinically healthy women, the majority of which were from the age group "41-50 years" (3,3±0,07%). The "medium risk group" was constituted from 95 women i.e.

26.4±0.02% and to the “low risk group” 253 women i.e. 70.3±0.03% respectively. Statically significant differences ( $p<0.001$ ). Belonging to high-risk groups increases the reliability of the pathological process in 2.2 (confidence interval 1.4-3.5) times ( $p<0.001$ ). Women of the high and medium risk group of breast cancer’s development based on the result of a comprehensive survey, a suspicion for a breast cancer was found at one woman (by histological examination - fibro adenoma), fibrocystic disease – at 18 women, fibro adenoma at 1 woman. In the low-risk group for breast cancer’s development the fibrocystic disease was detected for 4 women, nodal formations were not identified.

In order to conduct the verification of the work of the “Oncologist” program, as well as the developed by us algorithm based on the neural network, the additional group of patients (with FCD clinical records) was examined via the questionnaire method. All women were clinically examined beforehand; the diagnosis is confirmed by mammography and ultrasound (women after 35 years). Every participant passed the mandatory general clinical examination, thyroid ultrasound and laboratory testing of thyroid function. The program distributed all women to groups of risk after the introduction of survey’s data of women with FCD. Among the 360 examined groups of FCD patients, 213 women were assigned to the high risk group - 59,2±0,04%. 134 women were distributed to the “group of medium risk” 37,2±0,08%, to the “low rate risk group” 13 women 3.6±0.09% respectively. The differences are statically significant ( $p<0.001$ ). Affiliation to the high risk group increases the reliability of the pathological process’s development in 2.2 (confidence interval 1.4-3.5) times ( $p<0.001$ ). 7.7% i.e. 28 women with nodular form of FCD were distributed by the program to the high risk group. The sensitivity of the program is equal to 79.8%, specificity 79.0%. The program’s distribution of the women’s groups with FCD into risk groups confirms that the use of the artificial neural network model along with a developed computer program can reliably form the prognostic “risk group” for breast cancer’s development.

## Conclusions.

- The statistically significant of the risk factors were determined ( $p < 0.001$ ) for each age group in the Bucovina region, which makes up the specific weight. Therefore, from 1016 features of risk factors, which were determined at 297 of women, namely in percentage: age - 98%, premenstrual mammalgia - 89%, fibrocystic breast changes - 88%, chronic inflammation of the appendages - 73%, concomitant thyroid disease (endemic goiter) - 69%
- Automated system developed (based on the neural network), which allows to create groups of the increased risk of the pre-tumor diseases' occurrence, as well as the emergence of the breast cancer.
- Materials of the paper in the form of an automated system for rational diagnosis of breast diseases ensuring is introduced to the work of the oncology service in Chernivtsi region. In addition to the main mode, the program runs in the remote access mode, which means that a woman can take a questionnaire from home using a computer or mobile device.

## References

1. Forouzanfar M.H., Foreman K.J., Delossantos A.M. et al. (2011) Breast and cervical cancer in 187 countries between 1980 and 2010: a systematic analysis. *The Lancet*, 378 (9801): 1461–1484.
2. Чиссов В.И., Старинский В.В., Петрова Г.В. (2012) Злокачественные новообразования в России в 2010 году. «МНИОИ им. П.А. Герцена», Москва: 4–12.
3. Федоренко З.П., Михайлович Ю.Й., Гулак Л.О та ін. (2013) Рак в Україні, 2011–2012. Захворюваність, смертність, показники діяльності онкологічної служби. Бюл. Нац. канцер-реєстру України, 14. Національний інститут раку, Київ. 118 с.
4. Гилязутдинов И.А., Хасанов Р.Ш. (2004) Опухоли гормонально-зависимых и гормонопродуцирующих органов. «МЕДпресс-информ», Москва. 464 с.



5. Смоланка И.И., Складар С.Ю. (2008) Современный взгляд на проблему диагностики и лечения рака молочной железы. *Жіночий лікар*, 5: 8.
6. Семиглазов В.В., Топузов Э.Э. (2009) Рак молочной железы. «МЕДпресс-информ», Москва. 176 с.
7. Седаков И.Е., Александров А.И., Денисенко Я.В. и др. (2008) Ответ при лечении при мастопатии у женщин в менопаузе. *Онкология*, 4(38): 429–431.
8. Семиглазов В.В., Семиглазов В.Ф., Ермаченкова А.М. (2011) Минимальные формы рака молочной железы. *Вопросы онкологии*, Санкт-Петербург, 6(57): 702–706.
9. Peng S.Y., Wu K.C., Wang J.J. et al. (2007) Predicting post operative nausea and vomiting with the application of an artificial neural network. *Br. J. Anaesth.*, 98: 60–65.