"West Kazakhstan Medical University named after Marat Ospanov"

Annotation to the dissertation
for the Degree of Doctor of Philosophy (PhD)

"Breast Cancer Epidemiology, Genotyping, and Evaluation of Chemotherapy Response using DNA Double-Strand Breaks (γ-H2AX) Monitoring in Lymphocytes"

Specialty: 6D110100 – "Medicine"

AITMAGAMBETOVA MARZHAN ALTYNBEKOVNA

Scientific Advisor:
Assoc. Prof. Gaziza A. Smagulova, Head of the Pharmacology Department, West Kazakhstan Medical University named after Marat Ospanov, PhD in Medical Sciences.

Scientific Advisor:
Prof. Erbol Z. Bekmukhambetov, Chairman of the Board, National Healthcare Chamber, Doctor of Medical Sciences.

Foreign scientific Advisor:
MD PhD Irakli Kokhreidze, Vice-Rector for Clinical Work, Head of the Oncology Department at the First University Clinic, Tbilisi State Medical University.

Republic of Kazakhstan
Aktobe, 2023

ANNOTATION

of the Dissertation by Aitmagambetova Marzhan Altynbekovna on the topic: **"Breast Cancer Epidemiology, Genotyping, and Evaluation of Chemotherapy Response using DNA Double-Strand Breaks (γ-H2AX) Monitoring in Lymphocytes,"** submitted for the Degree of Doctor of Philosophy (PhD) in the specialty 6D110100 – "Medicine."

Scientific Advisor: Assoc. Prof. Gaziza A. Smagulova, PhD in Medical Sciences.
Scientific Advisor: Prof. Erbol Z. Bekmukhambetov, Doctor of Medical Sciences.
Foreign scientific Advisor: Prof. Irakli Kokhreidze, Doctor of Medical Sciences.

INTRODUCTION

Breast cancer (BC) is one of the most common malignancies among women worldwide and a leading cause of cancer-related deaths [Ginsburg O., et al. 2017]. When comparing data from GLOBOCAN 2018 and GLOBOCAN 2020, an increase in both the incidence and mortality rates of BC was noted. In 2020, there were 19.1 million new cases and 9.9 million deaths from BC worldwide [https://gco.iarc.fr]. The standardized rates for BC incidence and mortality worldwide in 2020 were 47.8 and 13.6 cases per 100,000 population, respectively [https://gco.iarc.fr]. According to the IARC Breast Source: Globocan 2020 data for the Republic of Kazakhstan, there were 35,366 newly diagnosed cases of malignancies, with 4,390 (12.4%) of them being BC cases. The total mortality from all malignancies was 20,959, with 1,654 thousand deaths (7.9%) attributed to breast cancer [https://gco.iarc.fr]. Based on the "Analytical Review of the Oncological Service Status in Aktobe Region for the year 2022," breast cancer ranks first in the structure of morbidity, accounting for 12.6%.

According to the Global CONCORD Program for Surveillance of Survival in Breast Cancer patients, the five-year survival rate was approximately 90% in the United States and Australia, and 40% in South Africa. In Asia, high five-year survival rates were observed in Israel (88%) and Japan (85%), while the lowest was in India (66.1%) [Allemani C., et al. 2018].

Medical-genetic counseling, followed by genetic testing, is conducted to identify hereditary predisposition to breast cancer, which is essential for verifying the genetic diagnosis in at-risk groups. This leads to personalized diagnostic, prognostic, and preventive approaches in treatment.

Breast cancer is a complex disease influenced by both genetic and non-genetic factors and, in most cases, without a clear inheritance pattern [Roberts E., et al. 2023]. It is known that 40% of malignancies with different anatomical localizations have a hereditary etiology, but for breast cancer, this proportion is 10%. Genes associated with breast cancer include BRCA1, BRCA2, CHEK2, NBS1, p53, ATM, among others. Detected hereditary mutations increase the risk of developing cancer [Laptiev S.A., et al. 2018]. According to epidemiological studies, only 15% - 20% of familial breast cancer cases are significantly linked to BRCA1 and BRCA2 mutations, while the

remaining 80% - 85% of familial risk are associated with other known and unknown gene polymorphisms [De Silva S., et al. 2019].

Studies dedicated to identifying associations of single nucleotide polymorphisms and mutations in other genes (besides BRCA1 and BRCA2) with the risk of developing breast cancer in the Kazakh population in Western Kazakhstan have not been conducted, and such research certainly holds scientific and practical interest. In the Republic of Kazakhstan, there is no established unified and clear algorithm for conducting molecular-genetic diagnostics. Investigating molecular-genetic markers associated with the risk of breast cancer development in the Kazakh population will help determine hereditary predisposition and facilitate early primary and secondary prevention.

Current research on cancer biomarkers is mainly focused on tumor-specific antigens. Recent reviews suggest using markers of cell proliferation, oncogene-induced aging, telomerase, DNA damage repair, and corresponding epigenetic markers as general cancer biomarkers, which are powerful and promising for cancer prediction, prognosis, therapy, and possibly cancer prevention [Shakyawar SK., et al. 2023].

The effectiveness of chemotherapy primarily relies on the destruction of rapidly dividing cancer cells by inducing various types of DNA damage. Double-strand breaks (DSBs) in DNA are the most dangerous type of DNA damage. DNA damage caused by various factors such as ionizing radiation, hypoxia, reactive oxygen species (ROS), certain chemicals, as well as errors in replication or transcription, activates processes for repairing these damages [Penninckx S., et al. 2021]. One of the initial events upon DNA damage repair activation is the phosphorylation of histone H2AX at serine 139 on each side of the break, forming γ-H2AX foci [Porcedda P., et al. 2009]. Currently, the analysis of phosphorylated histone protein H2AX (γH2AX) foci is the most sensitive method for detecting DNA double-strand breaks (DSBs). Immunofluorescent staining with anti-γH2AX antibodies allows visualization of nuclear foci and correlates with the number of DSBs [Raavi V., et al. 2020]. Phosphorylated histone protein γ-H2AX serves as a specific biomarker for cellular stress, particularly in the diagnosis and monitoring of neoplastic diseases.

Fluorescent microscopic examination is used for rapid and standardized determination of γ-H2AX, enabling the urgent assessment of DNA damage in clinical practice. The AKLIDES platform (Medipan, Dalevitz, Germany) allows complete automation of screening for antinuclear immunofluorescent antibodies [Bizzaro N., et al. 2018] and computational analysis of γ-H2AX foci, which has been successfully confirmed by several independent research groups [Willitzki A., et al. 2013]. This makes the detection of γ-H2AX an attractive biomarker that can serve as an early indicator of cancer.

Based on the above, the research aims and objectives have been formulated.

The objective of the research:

To study the epidemiology of breast cancer, gene polymorphisms associated with breast cancer, and to determine the value of monitoring γ-H2AX foci as a potential biomarker for chemotherapy response.

Research tasks:

1. Conduct an analysis of breast cancer incidence in Aktobe Region for the years 2014-2018 and 2022.
2. Investigate BRCA1 and BRCA2 gene polymorphisms in breast cancer patients using polymerase chain reaction (PCR) and identify single nucleotide polymorphisms in genes associated with the risk of breast cancer using Next-generation sequencing (NGS) with bioinformatics analysis.
3. Conduct a pilot study to assess the response to chemotherapy in breast cancer patients by monitoring DNA double-strand breaks in lymphocytes (γ-H2AX).

Scientific novelty:

1. A retrospective epidemiological analysis for 5 years (2014-2018 and 2022) was conducted using trend analysis and survival analysis.
2. For the first time, genetic analysis of BRCA1 (mutations 185delAG, 4153delA, 5382insC, 3819delGTAAA, 3875delGTCT, 300T>G (Cys61Gly), 2080delA) and BRCA2 (mutation 6174delT) genes was performed using PCR, and sequencing of 113 candidate genes was carried out using Next-generation sequencing (NGS) technology (Illumina). Molecular-genetic genotyping was also performed using chip-based systems and bioinformatics analysis in Kazakh women with breast cancer in Aktobe Region.
3. A pilot study of γ-H2AX foci as a biomarker for breast cancer chemotherapy response was conducted by monitoring DNA double-strand breaks in lymphocytes using the AKLIDES system (MEDIPAN, Germany), consisting of a fluorescent analyzer and AKLIDES Nuk software.

Theoretical and practical significance of the research:

The results of the epidemiological analysis expand our understanding of breast cancer incidence, 5-year survival rates, and dynamics analysis in Aktobe Region.

The genetic testing results for Kazakh women in Aktobe Region allow the identification of predisposition to hereditary forms of cancer and direct efforts towards cancer prevention and early diagnosis. It is recommended to include the determination of gene polymorphisms such as Rs137852985 (BRIP1), Rs2229774 (RARG), Rs2981582 (FGFR2), and Rs889312 (MAP3K1) in the screening program for breast cancer in at-risk groups (with a family history of cancer, young age) in Aktobe Region.

The implementation of monitoring γ-H2AX foci analysis (focilint mean - mean intensity value for all foci) in breast cancer patients is part of the personalized approach in the chemotherapy department of the Medical Center of ZKMU named after M. Ospanov.

Positions proposed for defense:

1. The analysis of dynamics analysis in newly diagnosed cases and overall incidence of breast cancer revealed a significant increase in incidence (R^2=0.3955, $p<0.021$) and overall incidence (R^2=0.9188, $p<0.001$) in Aktobe Region. The projected number of newly diagnosed cases of breast cancer in 2025 is estimated to be 218, and the projected overall incidence in 2025 is expected to rise to 1766 cases.
2. In Aktobe Region, BRCA1 gene polymorphism (5382insC) was detected in 0.72% and 300T>G (Cys61Gly) in 0.36% of women with breast cancer using PCR method.
   - Seven statistically significant risk-associated polymorphisms were identified, cataloged in GWAS and associated with the risk of developing breast cancer: RARG (Rs2229774), FGFR2 (Rs2981582), ATM (Rs1800057), MAP3K1 (Rs889312), BRCA2 (Rs11571833), FGFR2 (Rs7895676), FGFR2 (Rs1219648).
   - A high genotype-phenotype correlation was established between five single nucleotide polymorphisms and the risk of breast cancer development using five genetic models: rs2981582 of FGFR2 gene, rs2229774 of RARG gene, rs889312 of MAP3K1 gene, rs137852985 of BRIP1 gene, and rs137852576 of AR gene.
   - Top 32 prognostic risk factors for breast cancer development were determined, with risk level enhancements ranging from 69.7% to 90.6%, with the most important being: "Rs137852985 (BRIP1)", "Rs2229774 (RARG)", and "Rs2981582 (FGFR2)".
   - A prognostic risk model for breast cancer development was created with high model quality (0.88) and a 95% risk assessment when combining polymorphisms cataloged in GWAS: "Rs2229774 (RARG)", "Rs889312 (MAP3K1)", and "Age, years <54.0".
3. Statistically significant changes in the parameters of FITC breaks and APC repair of γH2AX foci in lymphocytes in women with breast cancer and benign diseases were revealed, which allows using γH2AX foci as a diagnostic marker for detecting breast cancer. The revealed statistically significant changes in the parameters of FITC breaks and APC repair of γH2AX foci in lymphocytes in women with breast cancer and benign diseases make it possible to use γH2AX foci as a diagnostic marker for detecting breast cancer. The prognostic model for evaluating the analysis of γ-H2AX foci as a marker for the diagnosis of breast cancer showed “good” and “average” quality of the forecast for the indicator “Mean intensity value for all foci” for the channels of FITC breaks (0.70) and APC repair (0.69).

- Statistically significant results allow the analysis of γ-H2AX foci in lymphocytes of breast cancer patients to be used as a possible early indicator of the effectiveness of chemotherapy.

The research results were presented and discussed at various scientific conferences:

-VI International Scientific Conference of Young Scientists and Students "Prospects for the Development of Biology, Medicine, and Pharmacy," December 7-8, 2018, Shymkent, Kazakhstan.
-VIII Annual International Scientific-Practical Conference "Current Issues in Medicine" and "Satellite Forum on Public Health and Healthcare Policy," April 10-12, 2019, Baku, Azerbaijan.
-III International Scientific-Educational Conference "Internationalization of Continuous Medical Education. A Look into the Future," April 25-26, 2019, Aktobe, Kazakhstan.
-VII Congress of Oncologists and Radiologists of Kazakhstan with international participation, Nur-Sultan, October 2019.
-12th Breast-Gynecological & Immunooncology International Cancer Conference, January 9-10, 2020, Cairo, Egypt.

Publications related to the dissertation:

12 scientific papers have been published, including 3 articles in Scopus-indexed journals: Reports of Practical Oncology and Radiotherapy (Cite Score percentile 32, SJR 0.337 in 2021); Carcinogenesis: Integrative Cancer Research (Cite Score percentile 72, SJR 1.178 in 2022, Q1-Cancer Research, WoS IF- 5.356, Q2); European Review for Medical and Pharmacological Sciences Radiotherapy (Cite Score percentile 78, SJR 0.634 in 2022), 3 articles in journals recommended by the Committee for Control in Education and Science of the Republic of Kazakhstan, 6 conference abstracts in international conference proceedings, including 2 indexed in Scopus and 1 indexed in Web of Science.

The dissertation research was carried out within the framework of funded scientific projects:

1. “Comparative experimental and clinical assessment and methods for correcting complications caused by chemotherapy of neoplastic processes in the mammary glands associated with mutations of the BRCA1, BRCA2 genes (experimental and clinical study” (State registration number No.: 0118RK01065, head – Zheksenova A.N., ZKMU named after M. Ospanova). MES RK, 2018-2020.
2. “New molecular genetic methods of pre-symptomatic diagnosis and methods of treatment of a number of significant diseases” (State registration number 0117RK00036, head - Ramazanova B. A., Asfendiyarov KazNMU). MES RK, 2019
3. Intra-university NTP grant “The role of microelement status, DNA damage in the development of oncopathology (using the example of the Aktobe region)”, head - Batyrova G.A., West Kazakhstan Medical University named after M. Ospanov 2020.

Results of the implementation of the research in practical healthcare and educational process:

- The research results have been integrated into the internship program at the Department of Oncology, specializing in "General Medicine," within the discipline "Oncology in Primary Care" for the 6th year and "Introduction to Oncology Patients in Primary Medical and Social Care" for the 7th year at Marat Ospanov West Kazakhstan

Medical University. The fundamental aspects of the dissertation work are employed in lecture materials and during practical sessions (see Appendix G).
- Intellectual property – 2 copyright certificates (Appendix D).

Author's personal contribution

Development of the goals and objectives of scientific research, collection and coordination of data from clinical, laboratory, molecular studies, statistical processing of research results, formulation of conclusions and practical recommendations.

Research Funding:

Dissertation Volume and Structure: The dissertation comprises 141 pages of computer text and includes the following sections: introduction, literature review, materials and methods of research, research findings, discussion of the obtained results, conclusion, summary, practical recommendations, reference list, and appendices. The bibliography encompasses 248 sources, including 241 international and 7 domestic scientific literature references. The doctoral dissertation is supported by 19 tables and 30 figures, which illustrate the essence of the conducted research.

Materials and Methods of Research

General Overview

This study was conducted at the Medical Center of M. Ospanov West Kazakhstan Medical University and was carried out within the framework of three funded projects: 1) "Comparative experimental-clinical assessment and methods of correction of complications caused by chemotherapy of neoplastic processes of the mammary glands associated with mutations in BRCA1, BRCA2 genes (experimental-clinical study)" (Registration number: 0118RK01065, Supervisor - Zhaksenova A.N., M. Ospanov WKMU). Ministry of Education and Science of the Republic of Kazakhstan, 2018-2020; 2) "New molecular-genetic methods of pre-symptomatic diagnostics and methods of treatment of certain significant diseases" (Registration number: 0117RK00036, Supervisor - Ramazanova B. A., Asfendiyarov Kazakh National Medical University). Ministry of Education and Science of the Republic of Kazakhstan, 2019; 3) "The role of microelement status, DNA damage in the development of oncopathology (based on the example of the Aktobe region)" Supervisor - Batyrova G.A., M. Ospanov WKMU, 2020.

The study consists of three main tasks, including epidemiology of breast cancer, genetic research, and examination of double-strand DNA breaks and repair in lymphocytes.

Epidemiological Analysis: In this part of the study, the incidence of breast cancer in the Aktobe region was analyzed for the years 2014-2018 and 2022. Data on new cases of the disease and overall incidence were used to identify trends and make forecasts for future years.

Genetic Research: This stage involved genetic testing of breast cancer patients using polymerase chain reaction (PCR) methods to detect mutations in the BRCA1 and BRCA2 genes. Next-generation sequencing (NGS) technology was also used for sequencing 113 candidate genes and conducting bioinformatic analysis.

Examination of DNA Breaks and Repair: At this stage, a pilot study was conducted to assess the response to chemotherapy in breast cancer patients by monitoring double-strand DNA breaks in lymphocytes using the AKLIDES system (MEDIPAN, Germany). Analysis of γ-H2AX (phosphorylated histone protein H2AX) foci was used as a predictor of chemotherapy effectiveness.

The study consists of three main tasks: epidemiology of breast cancer, genetic research, and examination of double-strand DNA breaks and repair in lymphocytes.

To address the first task of the study, information was obtained from the Electronic Registry of Oncological Patients in Aktobe Region at the Medical Center of M. Ospanov West Kazakhstan Medical University.

For the second task, genetic research was conducted at the Scientific Research Center of M. Ospanov West Kazakhstan Medical University in Aktobe and at the laboratory of the National Center for Biotechnology of the Ministry of Education and Science of the Republic of Kazakhstan in Astana.

The third task was carried out at the Scientific Research Center of M. Ospanov West Kazakhstan Medical University using the AKLIDES platform.

In accordance with the aims and objectives of the dissertation work, a research design was developed, defining the selection of materials and research methods. The research project was approved by the Bioethical Committee of M. Ospanov West Kazakhstan Medical University (Protocol No. 24 dated October 3, 2017).

Before participation in the study and after explaining the procedures, each participant signed a written informed consent to participate in the research.

Inclusion criteria: All female patients with breast cancer at stage I, II, IIIa, and IIIb, who have indications for surgical treatment and chemotherapy, and are aged above 18 years.

Within the scope of task 2 (Stage B), 149 ethnic Kazakh women were included for sequencing and bioinformatic analysis.

Exclusion criteria: Pregnant women, patients from clinical group 4 with severe comorbidities, and patients who refused to participate.

Task 1: Epidemiology of Breast Cancer in the Aktobe Region during 2014-2018 and 2020.: A retrospective epidemiological analysis was conducted over a span of 5 years (2014-2018 and 2022), utilizing disease incidence. To accomplish this task, information was extracted from the Electronic Cancer Patient Registry (ECPR), which included 30 forms and 7 templates. Population data for calculating breast cancer incidence rates in the Aktobe region were sourced from the Aktobe Regional Statistical Office (2014-2018 and 2022).

Task 2. Stage A: Determination of a standard diagnostic panel of polymorphisms in BRCA1 and BRCA2 genes in breast cancer patients using the polymerase chain reaction (PCR) method.

Research Design: A cross-sectional study was conducted, and genotyping of 278 breast cancer patients was performed at the Medical Center of Marat Ospanov West Kazakhstan Medical University (MCO WKMU). The patients had a confirmed

diagnosis of breast cancer, and the study used a standard diagnostic panel of 8 polymorphisms. The study period was 2 years (2018-2019). Allelic variants of BRCA1 (mutations 185delAG, 4153delA, 5382insC, 3819delGTAAA, 3875delGTCT, 300T>G (Cys61Gly), 2080delA) and BRCA2 (mutation 6174delT) genes were determined using the "Oncogenetics" reagent kit on human DNA extracted from peripheral blood. The identification of gene variants associated with the risk of oncopathology was performed using real-time polymerase chain reaction (PCR) method with the "GS-Genetics. OncoGenetics BRCA" probe. The research was carried out at the scientific molecular-genetic laboratory of the National Center for Research and Practical Oncology, Marat Ospanov West Kazakhstan Medical University. Approval from the local bioethics committee was obtained (approval number: № 20 dated 11.09.2017). Written informed consent was obtained from all study participants.

Task 2. Stage B: Next Generation Sequencing (NGS) and Bioinformatic Analysis of Genes in Breast Cancer Patients.

Research Design: This stage of the study followed a case-control design. The candidate gene sequencing using Next Generation Sequencing (NGS) technology (Illumina) and molecular-genetic genotyping using chip technology were conducted at the laboratory of the National Center for Biotechnology, Ministry of Education and Science of the Republic of Kazakhstan, in Astana. The bioinformatic analysis was performed at the Karaganda Medical University, under the supervision of Ph.D. Babenko D.B., the director of the Research Center.

The biomaterial genotyping (NGS sequencing) was performed using a panel of 113 single-nucleotide polymorphisms (SNPs) associated with the prognosis and progression of breast cancer in the Kazakh ethnic population. The study included the formation of patient groups based on inclusion criteria for collecting biological material: 149 patients of Kazakh ethnicity with morphologically confirmed breast cancer and 150 women in the control group. Peripheral venous blood was collected from the patients and individuals in the control group (conditionally healthy women without a history of oncological disease). DNA was extracted from the biological material, and a laboratory analysis was conducted to detect single-nucleotide polymorphisms in the genes in the biological samples.

Task 3. Analysis of γ-H2AX Foci as a Biomarker of Response to Chemotherapy in Breast Cancer Patients by Monitoring DNA Double-Strand Breaks in Lymphocytes using the "AKLIDES" Platform.

Research Design: This stage of the study followed a prospective cohort design and was conducted on two groups of patients. The main group consisted of 29 patients with a primary verified diagnosis of breast cancer (mean age 56, 95% CI: 51.4-60.76). The control group included 24 patients (mean age 43, 95% CI: 38.81-47.35) with histologically verified benign breast lesions. The study was conducted at the Medical Center (MC) and National Research Center (NRC) of the Marat Ospanov West Kazakhstan Medical University. The study material was peripheral venous blood, collected in a 10 ml EDTA tube, containing mononuclear cells. The analysis of γ-H2AX

foci in lymphocytes was performed using the "AKLIDES Nuk Human Lymphocyte Complete" immunofluorescence staining kit by Medipan.

The project was approved by the local bioethics committee (protocol No. 57, January 17, 2020). Written informed consent was obtained from all study subjects.

In the available literature, there were no recommendations found for conducting research on DNA breaks as a marker of chemotherapy efficacy in breast cancer. The proposed methodology represents a pilot project developed based on the recommendations of the AKLIDES platform developers from Germany. Considering the technical complexities of implementation (analysis within 1 hour after blood sampling, involvement of both inpatient and outpatient participants), a 4-stage scheme for conducting the study in the main group was proposed:

Before starting the 1st course of chemotherapy - Stage 1
After completing the 1st course of chemotherapy - Stage 2
Before starting the 2nd course of chemotherapy - Stage 3
Before starting the 3rd course of chemotherapy - Stage 4

In the control group, the analysis was conducted on women with histologically verified benign breast lesions. The patients in the main group underwent radical mastectomy by Letiagin, polychemotherapy, and radiation therapy on the TruBeam apparatus.

Methods of Statistical Data Analysis:

Task 1: The statistical analysis involved calculating frequencies and percentages for the studied variables according to the normality of the distribution. Confidence intervals (95% CI) were calculated, and overall survival was assessed using the Kaplan-Meier method. A significance level of $p < 0.05$ was considered statistically significant. Incidence trends were determined by the least squares method. Average annual growth rate (AGR), percentage changes were estimated using linear regression analysis, including a forecast index for 2025.

Task 2, Stage A: Descriptive statistics using SPSS were applied to calculate all indicators with a 95% confidence interval for data compilation.

Task 2, Stage B: The statistical analysis for bioinformatic analysis included the calculation of Hardy-Weinberg equilibrium, association analysis based on generalized linear models, and analysis of associations between polymorphisms and the phenotype of dominant genetic inheritance. The genotype-phenotype association was assessed using five different inheritance models: dominant, co-dominant, recessive, overdominant, and logarithmic models. To identify alleles significantly associated with the course of breast cancer in the Kazakh population (phenotype-genotype associations in the breast cancer group using the dominant inheritance model), logistic regression with Bonferroni correction for p-values in multiple comparisons was used.

Task 3: For describing numerical scales, mean and standard deviation were used in the form of "M ± S." All graphs for numerical variables depicted the arithmetic mean as a dot, the median as a horizontal line segment, the interquartile range as a rectangle, and the minimum and maximum levels as vertical line segments.

Comparisons between two groups for numerical variables were performed using the non-parametric Mann-Whitney test. The statistical significance of group differences for binary and categorical variables was determined using the Pearson chi-square test.

The analysis of variable dynamics, in the case of comparing two periods, was performed using the non-parametric Wilcoxon test, and for comparisons of three or more measurements, the non-parametric Friedman test was used.

The significance level was set at 0.05. Statistical data analysis was conducted using the Statistica 10 and SAS JMP 11 software packages.

Results of own research:

Epidemiology of breast cancer in Aktobe region: During the analysis of the Electronic Cancer Patient Registry (ECPR) from 2014 to 2018, a total of 891 new cases of breast cancer were registered in the Aktobe region. In the year 2022, there were 192 new cases of breast cancer reported. The five-year survival rate over the study period was 51.6% (95% CI: 50.45-52.89). In 2022, the five-year survival rate was 55.09%. The breast cancer mortality rate in 2022 according to ECPR data was 41 cases.

Dynamics analysis of newly diagnosed breast cancer cases showed an increase in incidence ($p<0.021$). The highest numbers were observed in 2019 (247 cases), possibly due to the effective implementation of the 'Densaulik' Health Development Program of the Republic of Kazakhstan for the years 2016-2019. A subsequent decrease in cases was noted in 2020 (155 cases), likely attributed to the COVID-19 pandemic when cancer screening programs were halted. The projected increase in incidence for 2025 is estimated to be 218 cases.

Overall breast cancer incidence dynamics analysis showed a growth in incidence ($p<0.001$). The highest incidence was observed in 2021 with 1561 cases, and a projected increase to 1766 cases by 2025.

Prevalence of BRCA1 and BRCA2 gene polymorphisms in breast cancer patients: Out of the 278 women with breast cancer, a family history of breast cancer was identified in 33 cases (11.8%). Using the polymerase chain reaction (PCR) method, 3 cases of BRCA1 gene mutation were detected among the 278 patients. Two patients had a family history of breast cancer and were found to have the BRCA1 polymorphism 5382insC, while the third patient did not have a family history of breast cancer but had the BRCA1 polymorphism 300T>G (Cys61Gly).

Results of NGS sequencing and bioinformatic analysis of genes in breast cancer patients - Stage B: A total of 149 patients with verified diagnosis of breast cancer, who were admitted to the Marat Ospanov West Kazakhstan Medical University, were included in the study during 2018. The control group comprised 150 apparently healthy women enrolled in the study as part of the project of the Ministry of Education and Science of the Republic of Kazakhstan. All participants were unrelated Kazakh women based on their pedigree.

The genotyping panel included 113 polymorphisms located in various regions of different chromosomes, as well as in various functional regions of genes and intergenic regions according to GWAS (Genome-Wide Association Study) data.

Comparative evaluation of allele and genotype differences between breast cancer patients and the control group revealed 28 statistically significant polymorphisms associated with breast cancer. Among the identified polymorphisms showing statistically significant differences in the studied groups according to the GWAS catalog (https://www.ebi.ac.uk/gwas/), 7 risk-associated polymorphisms for breast cancer development were determined: RARG (Rs2229774), FGFR2 (Rs2981582), ATM (Rs1800057), MAP3K1 (Rs889312), BRCA2 (Rs11571833), FGFR2 (Rs7895676), FGFR2 (Rs1219648).

A genotype-phenotype association was determined for 28 single nucleotide polymorphisms (SNPs) and the risk of developing breast cancer using five genetic models (codominant, dominant, recessive, overdominant, and log-additive models). The analysis revealed the following associations:

Polymorphism rs2981582 of the FGFR2 gene increases the risk of developing breast cancer in four inheritance models.

Polymorphism rs2229774 of the RARG gene increases the risk of developing breast cancer in three inheritance models.

Polymorphism rs889312 of the MAP3K1 gene increases the risk of developing breast cancer in two inheritance models.

Polymorphism rs137852985 of the BRIP1 gene increases the risk of developing breast cancer in four inheritance models.

The A/G genotype of rs137852576 of the AR gene increases the risk of developing breast cancer by 2.24 times in the codominant model.

The results of the one-factor analysis of risk for the target variable "Breast Cancer" (BC) showed that 32 factors have a statistically significant influence on the target variable with risk levels ranging from 50.7% to 90.6%. Under the influence of factors from the TOP-40 list, the risks of developing "RMJ" increase from 47.6% to 100.0%. The leading statistically significant factors with increased risk levels from 69.7% to 90.6% are the polymorphisms: "Rs137852985 (TC, TT)", "Rs2229774 (AG, AA)", and "Rs2981582 (AG, AA)". The TOP-40 list is completed by the statistically significant factors "Rs1219648 (GG)", "Rs1800470 (TT, CC)", and "Rs12762549 (CG, GG)", which increase the risk level from 53.0% to 63.0%.

Based on the results of the decision tree method for the target variable "BC", 6 risk classes were identified with risk levels ranging from 0.0% to 100.0%, using the following five influencing factors: "Rs137852985", "Rs757229", "Rs2981582", "Age, years < 56.0", and "Age, years ≥ 49.0". The highly risky class with a risk level of 100.0% is determined based on the combination of factors "Rs137852985 (TC, TT)" and "Rs2981582 (AA, GG)". The predictive quality of the constructed model is of high level.

Results of the decision tree method for the target variable "Breast Cancer" (based on the GWAS catalog) revealed 6 risk classes with risk levels ranging from 1.9% to 95.8%, using the following five influencing factors: "Age, years", "Rs2229774", "Rs2981582", "Rs889312", and "Rs1800057". The high-risk class with

a risk level of 95.8% is determined based on the combination of factors "Age, years <54.0", "Rs2229774 (AG)", and "Rs889312 (AA, CC)". The predictive quality of the constructed model is of a high level.

Results of the analysis of γ-H2AX DNA foci as a biomarker for the response to chemotherapy in breast cancer:

The study was conducted on patients with breast cancer (main group) and women with benign breast diseases (control group) using the AKLIDES platform. The main group included 29 patients with breast cancer, with a mean age of 56.10 ± 12.23. Based on the disease stages: 25 (86.2%) patients had stage II, and 4 (13.8%) patients had stage III. Immunohistochemistry analysis of the tumors showed that 3 (10.3%) patients had Luminal A subtype, 21 (72.4%) had Luminal B subtype, 4 (13.8%) had Triple-negative subtype, and 1 (3.4%) had HER2-positive breast cancer.

The results of the analysis of γ-H2AX foci in the control and main groups (stage 1) for the break channel (FITC) and the repair channel (APC) are as follows: In the FITC break channel, three indicators showed statistically significant differences between the two compared groups. The "Average nucleus diameter" was higher in the main group compared to the control group ($p = 0.0382$), while the "Mean intensity value for all foci" was lower in the main group compared to the control group ($p = 0.0166$). The "Number of overlapping foci in both channels" was higher in the main group than in the control group ($p = 0.0486$).

When comparing the indicators in the main and control groups in the APC repair channel, two statistically significant indicators were found. The "Nuclei with increased intensity of fluorescence" were lower in the main group compared to the control group ($p = 0.0166$), and the "Mean intensity value for all foci" was lower in the main group compared to the control group ($p = 0.0118$).

The results of the statistical one-factor prediction of the target indicator "BC, +" for quantitative and binary factors are as follows: Twelve factors have statistical significance in influencing the risk of developing "RMC, +" with risk levels ranging from 69.0% to 100.0%. The leading statistically significant factors for the development of BC with risks from 75.0% to 81.0% are "Age, years ≥ 48.0", "Mean intensity value for all foci (Stage 1) < 341.6", and "Age groups (After 50 years)". The list of statistically significant factors concludes with "Percentage of nuclei with foci in clusters of low intensity (Stage 1) ≥ 87.3", "Nuclei with increased intensity of fluorescence (Stage 1) < 35.6", and "Average foci diameter (Stage 1) ≥ 0.5", which increase the risk level from 69.0% to 73.7%.

The prognostic evaluation of γ-H2AX foci analysis as a diagnostic marker was performed using ROC analysis. The model showed “good” and “average” predictive quality of the diagnostic marker based on the “Mean intensity value for all lesions” in the FITC (0.70) and APC repair channels (0.69), respectively ($p = 0.001$).

The analysis of the dynamics of indicators in the FITC channel and APC repair channel during the periods of chemotherapy ("Stage 1", "Stage 2", "Stage 3", and "Stage 4") revealed significant differences in two indicators in the FITC

channel. The "Mean intensity value for all foci" in the main group increased after receiving the 1st course of chemotherapy, with a gradual decrease in its value (+4.6%, -7%, -11.7%; $p = 0.0075$). The indicator "Number of overlapping foci in two channels" increased after the 1st course of chemotherapy (on average by 15.0), followed by a gradual decrease until the start of the 3rd course of chemotherapy ($p = 0.0237$). These findings suggest a significant response of the organism to the 1st course of chemotherapy.

During the considered period in the APC repair channel, nine indicators significantly changed: the number of cells with foci ($p = 0.0465$), the total number of foci ($p = 0.0007$), the number of detected clusters ($p = 0.0020$), the mean number of foci per cell ($p < 0.0001$), the mean value of all foci per cell ($p = 0.0003$), the percentage of damaged cells ($p < 0.0001$), the mean value of all foci in clusters with low intensity ($p < 0.0001$), the mean value of foci in clusters with low intensity in the cluster ($p = 0.0003$), and the percentage of nuclei with foci in clusters with low intensity ($p = 0.0021$). The statistically significant changes in these 11 indicators in the APC repair channel with specific shifts before the start of the 2nd course of chemotherapy indicate the organism's response to the conducted chemotherapy, manifested by the enhancement of the DNA repair process in patients with breast cancer.

The dynamics of the ratios of double-strand breaks (DSBs) to repairs were analyzed. It was found that all the indicators of γ-H2AX parameters in the APC repair channel were statistically significantly higher than those in the FITC channel.

In the course of the study, out of 29 patients, the outcome was unknown (dislocation) for two women, and four patients died. The three-year survival rate was 85%.

B conclusion, our study investigated the epidemiology of breast cancer, genotyping with PCR and NGS sequencing in breast cancer patients, and conducted a pilot study on the analysis of γH2AX foci as a biomarker for response to breast cancer chemotherapy by monitoring DNA double-strand breaks in lymphocytes on the AKLIDES platform.

Based on the obtained results, the following conclusions can be drawn:

1. The analysis of dynamics in newly diagnosed cases and overall incidence of breast cancer showed an increasing trend with $R^2=0.3955$ ($p<0.021$) and $R^2=0.9188$ ($p<0.001$), respectively. The projected number of newly diagnosed breast cancer cases in 2025 will be 218. The projected overall incidence in 2025 will rise to 1766 cases.
2. In the Aktobe region, the BRCA1 polymorphisms (5382insC and 300T>G, Cys61Gly) were detected in 0.72% and 0.36% of breast cancer patients, respectively, using PCR. NGS sequencing revealed 6 statistically significant polymorphisms in ATM (rs1800057), RARG (rs2229774), BRCA2 (rs11571833), MAP3K1 (rs889312), FGFR2 (rs2981582), and BRIP1 (Rs137852985) genes

associated with a high risk of breast cancer development in Kazakh women in the Aktobe region.

3. Statistically significant changes in the parameters of FITC breaks and APC repair of γH2AX foci in lymphocytes in women with breast cancer and benign diseases were revealed, which allows using γH2AX foci as a diagnostic marker for detecting breast cancer. The revealed statistically significant changes in the parameters of FITC breaks and APC repair of γH2AX foci in lymphocytes in women with breast cancer and benign diseases make it possible to use γH2AX foci as a diagnostic marker for detecting breast cancer. The prognostic model for evaluating the analysis of γ-H2AX foci as a marker for the diagnosis of breast cancer showed "good" and "average" quality of the forecast for the indicator "Mean intensity value for all foci" for the channels of FITC breaks (0.70) and APC repair (0.69). Statistically significant results allow the analysis of γ-H2AX foci in lymphocytes of breast cancer patients to be used as a possible early indicator of the effectiveness of chemotherapy.