

ПРОБЛЕМЫ СОЦИАЛЬНО ЗНАЧИМЫХ ЗАБОЛЕВАНИЙ

PROBLEMS OF SOCIALLY SIGNIFICANT DISEASES

© КОЛЛЕКТИВ АВТОРОВ, 2024

Ермолицкая М.З.^{1,2}, Кику П.Ф.³, Абакумов А.И.¹

Модель оценки влияния показателей качества жизни на заболеваемость раком молочной железы

¹ФГБУН «Институт автоматики и процессов управления ДВО РАН», 690041, Владивосток, Россия;

²ФГБОУ ВО «Владивостокский государственный университет», 690014, Владивосток, Россия;

³ФГАОУ ВО «Дальневосточный федеральный университет», 690922, Владивосток, о. Русский, п. Аякс, Россия

РЕЗЮМЕ

Введение. Злокачественные новообразования молочной железы являются ведущей онкологической патологией среди женского населения Приморского края. Выявление взаимосвязи между заболеваемостью и показателями качества жизни (КЖ) наряду с современными методами диагностики позволяет совершенствовать профилактические мероприятия по снижению заболеваемости на региональном уровне.

Цель исследования — разработка регрессионной модели, описывающей статистические взаимосвязи между заболеваемостью раком молочной железы и социально-экономическими показателями КЖ населения Приморского края.

Материалы и методы. Исходная выборка данных состояла из 17 показателей КЖ населения Приморского края за 1994–2020 гг. Для уменьшения размерности данных воспользовались методом главных компонент, для построения модели — регрессионным анализом. Качество построенной модели проверяли на основе расчёта коэффициента детерминации, среднеквадратической ошибки и ошибки аппроксимации.

Результаты. В исследовании выявлены 15 показателей КЖ населения, взаимосвязанные с заболеваемостью раком молочной железы на территории Приморского края. Анализ главных компонент позволил сгруппировать показатели КЖ в три главные компоненты. Первая компонента объясняет 80,8% общей дисперсии, вторая — 10%, третья — 4%. В первую компоненту вошли показатели, характеризующие социально-экономические условия жизни населения, во вторую — медико-социальные, в третью — показатели неравенства денежных доходов, характеризующие социальную дифференциацию общества. На основе выделенных главных компонент разработана регрессионная модель.

Ограничения исследования. Материалы исследования ограничены результатами статистического анализа 17 показателей КЖ населения Приморского края за 1994–2020 гг. и применением разработанной регрессионной модели на региональном уровне.

Заключение. Результаты данного исследования позволили выявить взаимосвязь заболеваемости раком молочной железы с факторами риска и разработать прогнозную модель, что может быть полезным при планировании профилактических мероприятий по улучшению КЖ и снижению заболеваемости на региональном уровне.

Ключевые слова: рак молочной железы; заболеваемость; качество жизни; факторы; модель

Соблюдение этических стандартов. Исследование не требует представления заключения комитета по биомедицинской этике.

Для цитирования: Ермолицкая М.З., Кику П.Ф., Абакумов А.И. Модель оценки влияния показателей качества жизни на заболеваемость раком молочной железы. *Здравоохранение Российской Федерации. Здравоохранение Российской Федерации.* 2024; 68(2): 116–122. <https://doi.org/10.47470/0044-197X-2024-68-2-116-122> <https://elibrary.ru/>

Для корреспонденции: Ермолицкая Марина Захаровна, канд. биол. наук, ст. науч. сотр. Института автоматики и процессов управления ДВО РАН, 690041, Владивосток. E-mail: ermmmz@mail.ru

Участие авторов: Ермолицкая М.З. — сбор и статистическая обработка данных, построение модели, написание текста, оформление рисунка, работа с литературой; Кику П.Ф. — концепция и дизайн исследования, написание текста; Абакумов А.И. — концепция и дизайн исследования, редактирование текста рукописи. Все соавторы утвердили окончательный вариант статьи и несут ответственность за целостность всех частей статьи.

Благодарность. Авторы выражают благодарность доктору мед. наук, проф. В.И. Апанасевичу за консультирование в процессе исследования.

Финансирование. Работа выполнена в рамках государственного задания № 0202-2022-0002, тема «Разработка передовых методов и технологий создания интеллектуальных информационных и управляющих систем».

Конфликт интересов. Авторы декларируют отсутствие явных и потенциальных конфликтов интересов в связи с публикацией данной статьи.

Marina Z. Ermolitskaya^{1,2}, Pavel F. Kiku³, Aleksandr I. Abakumov¹

Model for assessing the influence of quality of life indicators on the breast cancer incidence

¹Institute of Automation and Control Processes of the Far Eastern Branch of the Russian Academy of Sciences, Vladivostok, 690041, Russian Federation;

²Vladivostok State University, Vladivostok, 690014, Russian Federation;

³Far Eastern Federal University, School of Medicine, Vladivostok, 690922, Russian Federation

ABSTRACT

Introduction. Malignant neoplasms of the breast are the leading oncological pathology among the female population of the Primorsky region. Identification of the relationship between the incidence rate and quality of life indicators, along with modern diagnostic methods, makes it possible to improve preventive measures to reduce the prevalence in the population at the regional level.

Purpose. Development of a regression model that describes the impact of socio-economic indicators of quality of life on the incidence of breast cancer in the population of the Primorsky region.

Materials and methods. The initial data sample consisted of seventeen indicators of the quality of life in the population of Primorsky region for the period from 1994 to 2020. To reduce the dimensionality of the data, the principal component method was used, and regression analysis was used to build the model. The quality of the constructed model was checked on the base of the calculation of the coefficient of determination, the standard error, and the approximation error.

Results. There were identified 15 indicators of the quality of life in the population that significantly affect the pathology of breast cancer in the Primorsky region. Principal component analysis has made it possible to group the quality of life indicators into three major compartments. The first component explains 80.8% of the variance, the second — 10%, the third — 4%. The first compartment included indicators characterizing the socio-economic aspects of the life of the population, the second — medical and social, and the third — statistical indicators of inequality in monetary incomes of the population, characterizing the social differentiation of society. A regression model has been developed on the principal compartments.

Research limitations. The research materials are limited to the results of statistical analysis of 17 indicators of the quality of life of the population of the Primorsky region for the period from 1994 to 2020 and the application of the developed regression model at the regional level.

Conclusion. The results of this study made it possible to identify the relationship between the incidence of breast cancer and risk factors and develop a predictive model, which can be useful in planning preventive measures to improve the quality of life and reduce the incidence at the regional level.

Keywords: breast cancer; incidence; quality of life; factors; model

Compliance with ethical standards. The study does not require the submission of the conclusion of the biomedical ethics committee.

For citation: Ermolitskaya M.Z., Kiku P.F., Abakumov A.I. Model for assessing the influence of quality of life indicators on the breast cancer incidence. *Zdravookhranenie Rossiiskoi Federatsii / Health Care of the Russian Federation, Russian journal.* 2024; 68(2): 116–120. <https://doi.org/10.47470/0044-197X-2024-68-2-116-122> <https://elibrary.ru/abcdef> (in Russian)

For correspondence: Marina Z. Ermolitskaya, Ph.D., senior researcher, Institute of Automation and Control Processes, Far Eastern Branch of the Russian Academy of Sciences, Vladivostok, 690041, Russian Federation. E-mail: ermz@mail.ru

Contribution of the authors: Ermolitskaya M.Z. — collection and statistical processing of data, building a model, writing a text, working with literature; Kiku P.F. — research concept and design, writing the text; Abakumov A.I. — research concept and design, editing. All authors are responsible for the integrity of all parts of the manuscript and approval of its final version.

Gratitude. The authors are grateful to Professor Apanasevich V.I. for advice during the research.

Acknowledgment. The work was carried out within the framework of state assignment No. 0202-2022-0002 «Development of advanced methods and technologies for creating intelligent information and control systems».

Conflict of interest. The authors declare no conflict of interest.

Received: November 18, 2022 / Revised: February 10, 2023 / Accepted: March 02, 2023 / Published: April 29, 2024

Введение

Злокачественные новообразования молочной железы занимают лидирующее место среди онкологических заболеваний в мире. По данным Всемирной организации здравоохранения (ВОЗ), в 2020 г. в мире было диагностировано 2,3 млн новых случаев рака молочной железы (РМЖ) и зарегистрировано 685 тыс. случаев смерти от этой болезни. По числу смертей от рака женского населения РМЖ занимает 2-е место после рака лёгкого. Наиболее высокие показатели заболеваемости зарегистрированы в США и странах Западной Европы, низкие показатели — в Китае и Японии [1]. По оценкам специалистов Международного агентства по изучению рака, к 2040 г. число ежегодно впервые регистрируемых случаев онкологических заболеваний возрастёт на 47% и достигнет 28,4 млн. Увеличение показателей ожидается в странах с низким и средним индексом развития человеческого потенциала [2, 3].

В России заболеваемость РМЖ находится на 1-м месте в структуре заболеваемости злокачественными новообразованиями женского населения (в 2020 г. — 21,1%). Число пациентов, состоящих на учёте в онкологических учреждениях, постоянно растёт: в 1994 г. — 266 497 человек (179,9⁰/₀₀₀₀), в 2004 г. — 408 399 (391,7⁰/₀₀₀₀), в 2019 г. — 718 803 (490⁰/₀₀₀₀), в 2020 г. — 734 587 (502⁰/₀₀₀₀). Общий показатель заболеваемости в 2019 г. составил 50,8⁰/₀₀₀₀, стандартизованный показатель (СП) — 30,7⁰/₀₀₀₀ (среди женского населения — 94 и 53,9⁰/₀₀₀₀ соответственно). В 2020 г. наблюдалось снижение показателей заболеваемости РМЖ: общий показатель заболеваемости составил 47,5⁰/₀₀₀₀ (82,8⁰/₀₀₀₀ среди женщин), СП — 27⁰/₀₀₀₀ (47,4⁰/₀₀₀₀ среди женщин). Такое снижение обусловлено эпидемиологической обстановкой по COVID-19, из-за которой диспансеризация населения была приостановлена, и процесс выявления заболеваний был затруднён.



Динамика контингента больных раком молочной железы в Приморском крае, ‰.
Trend in the contingent of patients with breast cancer in the Primorsky region, ‰.

Смертность от РМЖ в России остаётся высокой. Абсолютное число умерших от этого заболевания в 2004 г. составило 22 757 человек, в 2019 г. — 21 871, в 2020 г. — 21 634. Наблюдается медленное снижение смертности в среднем на 0,2% в год, что связано с ранним выявлением и своевременным лечением больных [4].

Знания об уровне заболеваемости и смертности от РМЖ, а также учёт факторов, влияющих на их рост, имеют важное значение для общественного здравоохранения, особенно при выработке стратегии повышения качества жизни (КЖ) и укрепления здоровья населения. Проведены различные исследования как отечественных, так и зарубежных авторов, подтверждающие влияние факторов риска на заболеваемость РМЖ. К факторам риска относят наследственность, сопутствующие заболевания, социально-экономические, санитарно-гигиенические, экологические и другие факторы. Особое внимание уделяют образу жизни человека: питанию, потреблению алкоголя и табака, физической активности, семейному положению [5–7]. Исследованию подвергаются отдельные категории граждан (по возрасту, полу, национальной структуре) или всё население в целом с учётом территориально-географических особенностей мест проживания [8–13]. При этом используют современные методы статистического анализа и машинного обучения, что позволяет разрабатывать и модифицировать модели, которые могут быть использованы для диагностики и прогнозирования уровня заболеваемости РМЖ [14–21].

В Приморском крае до 2020 г. также наблюдалась тенденция возрастания заболеваемости РМЖ. В 2004 г. на учёте в онкологических учреждениях состояли 226,4‰ (29,6‰ — впервые выявленные случаи), в 2019 г. — 471,4‰ (44,6‰ — впервые выявленные случаи) (рисунки). В 2020 г. впервые зарегистрировано 750 (76,3‰) женщин с диагнозом РМЖ, что на 200 человек меньше, чем в 2019 г. (95,8‰). Это можно объяснить пандемией COVID-19, в период которой темпы постановки на учёт новых больных снизились. В последующие 1–2 года следует ожидать всплеск заболеваемости с увеличением доли запущенных случаев РМЖ.

Цель данного исследования — разработка регрессионной модели, описывающей влияние социально-экономических показателей КЖ на заболеваемость РМЖ населения Приморского края.

Материалы и методы

Исходная выборка данных по заболеваемости РМЖ состоит из 27 наблюдений за период с 1994 по 2020 г. и 17 показателей КЖ населения Приморского края в целом:

- доход (среднедушевые денежные доходы в месяц, руб.);
- заработная плата (среднемесячная номинальная начисленная заработная плата в расчёте на 1 работника);
- прожиточный уровень (величина прожиточного уровня в среднем на душу населения в месяц, руб.);
- инвестиции (инвестиции в основной капитал на душу населения в фактически действующих ценах, тыс. руб.);
- коэффициент Джини (индекс концентрации доходов);
- объём платных услуг населению в расчёте на душу населения (тыс. руб.);
- жильё (введено в действие общей площади жилых домов и общежитий, тыс. м²);
- коэффициент фондов (соотношение доходов 10% наиболее и 10% наименее обеспеченного населения);
- валовый региональный продукт на душу населения (тыс. руб.);
- численность населения (на 1 января, тысяч человек);
- естественный прирост на 1000 человек населения;
- численность врачей на 10 тыс. человек населения;
- численность среднего медицинского персонала на 10 тыс. человек населения;
- обеспеченность больничными койками на 10 тыс. населения;
- индекс человеческого развития;
- индекс образования;
- уровень безработицы (по методологии Международной организации труда).

Численные значения данных показателей были получены из материалов официального сайта Федеральной службы государственной статистики (Росстат) [3, 4]. Показатели заболеваемости согласно Международной классификации болезней (МКБ-10) по классу «Злокачественное новообразование молочной железы» (C50) на 100 тыс. населения предоставлены медико-информационным аналитическим центром Приморского края.

Данные собирали в программе MS Excel 2010, обработку и анализ исходных данных осуществляли в программе RStudio v. 1.0.153. Проверку нулевых гипотез проводили методом Р. Фишера на заданном уровне зна-

Таблица 1. Значимые коэффициенты корреляции Кендалла между заболеваемостью рака молочной железы (РМЖ) и показателями качества жизни (КЖ)

Table 1. Significant Kendall correlation coefficients between the incidence and quality of life indicators

Показатель Indicator	Коэффициент корреляции Correlation coefficient
Доход Income	0,903
Среднемесячная зарплата Average monthly salary	0,915
Прожиточный уровень Subsistence level	0,892
Инвестиции на душу населения Investment per capita	0,698
Коэффициент Джини Gini coefficient	0,693
Платные услуги населению Paid services to population	0,903
Жилье Housing	0,439
Коэффициент фондов Funds ratio	0,616
Валовый региональный продукт на душу населения Gross regional product per capita	0,915
Численность населения Population	–0,915
Численность среднего медицинского персонала Number of paramedical personnel	–0,469
Обеспеченность больничными койками Availability of hospital beds	–0,436
Индекс человеческого развития Human development Index	0,816
Индекс образования Education Index	0,831
Уровень безработицы Unemployment rate	–0,847

чимости 0,05 [22]. Для проверки на нормальность распределения исходных данных использовали критерий Шапиро–Уилка. Выявление показателей, существенно влияющих на заболеваемость РМЖ, осуществляли методом τ Кендалла. Для проверки значимости коэффициентов корреляции, значения которых меньше 0,5, применяли t -критерий Стьюдента.

В ходе исследования для уменьшения размерности данных без значительной потери информации воспользовались методом главных компонент (разложение Карунена–Лоева), где главные компоненты — это линейные комбинации наблюдаемых показателей с соответствующими весами. Оптимальную размерность пространства главных компонент определяли с помощью критерия Кайзера–Гутмана и правила Р. Кеттелла [23, 24]. Предварительно исходные данные были стандартизованы (преобразованы в Z-шкалу с нулевым средним и единичной дисперсией). Взаимосвязь между главными компонентами и заболеваемостью выявляли методом Спирмена.

Для построения прогнозной модели применили регрессионный анализ с главными компонентами в качестве независимых переменных. При этом использовали классические инструменты регрессионного анализа: F-критерий Фишера, t -критерий Стьюдента. Исходная выборка данных была поделена на обучающую и тестовую выборки в соотношении 3 к 1. Для определения качества построенной модели были рассчитаны коэффициент детерминации, среднеквадратическая ошибка, ошибка аппроксимации и значение критерия Акаике по всей выборке и отдельно — по обучающей и тестовой выборкам.

Результаты

Статистический анализ показал, что ряд данных по уровню заболеваемости не соответствует нормальному закону распределения ($p = 0,003$). Поэтому для выявления показателей, значимо ассоциированных с уровнем заболе-

ваемости РМЖ, использовали непараметрический метод, в результате которого были выявлены 15 показателей КЖ, значимо ассоциированных с заболеваемостью (табл. 1).

Применение анализа главных компонент позволило заменить 15 выделенных скоррелированных показателей меньшим числом нескоррелированных переменных, которые сохраняют наибольшее количество информации, сохранив в исходных данных. Было решено оставить три главные компоненты, объясняющие 95% общей дисперсии исходных данных. При этом первая компонента объясняет 80,8% дисперсии, вторая — 10%, третья — 4% (табл. 2).

В первую компоненту вошли показатели, характеризующие социально-экономические условия жизни населения, во вторую — медико-социальные показатели, в третью — статистические показатели неравенства денежных доходов населения, характеризующие социальную дифференциацию общества. По общности, сохранённой в факторной структуре, наибольшие значения имеют следующие показатели: обеспеченность больничными койками (0,388), инвестиции на душу населения (0,376), жильё (0,352) и численность среднего медицинского персонала (0,265).

Для построения прогнозной модели использовали регрессионный анализ на выделенных главных компонентах. Нулевую гипотезу об отсутствии линейной зависимости между количеством заболевших людей и главными компонентами отвергли ($p < 0,001$). Уравнение зависимости имеет следующий вид:

$$Incidence = 0,7104 \times PC_1 + 0,1424 \times PC_2 + 0,3263 \times PC_3,$$

где *Incidence* — заболеваемость; PC_1 , PC_2 , PC_3 — главные компоненты.

Все коэффициенты регрессии в полученном уравнении значимы и по критерию Стьюдента ($p \ll 0,05$), и по построенным доверительным интервалам: для PC_1 — 0,644–0,777; PC_2 — 0,049–0,236; PC_3 — 0,225–0,427.

Таблица 2. Результаты анализа главных компонент**Table 2.** Results of principal component analysis

Главные компоненты Principal compartments	Показатель Indicator	Факторные нагрузки Factor Loads	Общность* Component communalities*	Коэффициент для главных компонент Main compartment coefficient
PC ₁	Доход Income	0,276	0,141	0,251
	Среднемесячная зарплата Average monthly salary	0,276	0,134	0,238
	Прожиточный уровень Subsistence level	0,278	0,146	0,268
	Платные услуги населению Paid services to population	0,278	0,124	0,202
	Валовый региональный продукт на душу населения Gross regional product per capita	0,277	0,135	0,231
	Численность населения Population	−0,277	0,121	−0,184
	Индекс человеческого развития Human development index	0,278	0,107	0,108
PC ₂	Численность среднего медицинского персонала Number of paramedical personnel	0,459	0,265	0,362
	Обеспеченность больничными койками Availability of hospital beds	0,595	0,388	0,437
PC ₃	Инвестиции на душу населения Investment per capita	−0,544	0,376	−0,641
	Коэффициент Джини Gini coefficient	−0,244	0,141	−0,234
	Жилье Housing	−0,421	0,352	−0,568
	Коэффициент фондов Funds ratio	−0,344	0,205	−0,341
	Индекс образования Education index	0,319	0,175	0,303
	Уровень безработицы Unemployment rate	−0,269	0,142	−0,271

Примечание. *Общность измеряемой переменной (сумма квадратов факторных нагрузок по строке ($\sum PC_i^2$)) интерпретируется как доля информации переменной, сохранённой в факторной структуре.

Note. * The commonality of the measured variable (sum of squared factor loadings per row ($\sum PC_i^2$)) is interpreted as the proportion of the variable's information stored in the factor structure.

Таблица 3. Оценки качества регрессионной модели**Table 3.** Regression Model Quality Estimates

Выборка Sampling	Коэффициент детерминации Determination coefficient	Среднеквадратическая ошибка RMS error	Ошибка аппроксимации Approximation error	Значение критерия Акаике AIC
По всей выборке данных Across the entire data set	0,97	0,028	0,018	11,55
На обучающей выборке On the training set	0,88	0,006	0,107	5,12
На тестовой выборке On the test sample	0,99	0,003	0,037	10,73

Согласно коэффициентам регрессии наибольшее влияние на заболеваемость РМЖ оказывают показатели, сгруппированные в первой компоненте. При этом коэффициент корреляции *Incidence* с PC_1 составил 0,97, с PC_2 — 0,53, с PC_3 — 0,63.

Для оценки качества регрессионной модели исследуемую выборку поделили на обучающую и тестовую в отношении 3 к 1, получили следующие результаты (табл. 3).

Обсуждение

На заболеваемость РМЖ, кроме наследственных и генетических факторов, оказывают влияние и группы показателей, характеризующие уровень жизни и КЖ населения, что отражено в ряде работ российских и зарубежных авторов [5, 7, 8, 11, 12]. К наиболее распространённым факторам риска относят возраст, образ жизни (физическая активность, питание, потребление алкоголя и табака),

уровень образования, национальность, сопутствующие заболевания и факторы окружающей среды. Результаты данного исследования подтверждают наличие взаимосвязи заболеваемости РМЖ с социально-экономическими показателями на территории Приморского края, включая показатели материального благополучия населения, с учётом социальной дифференциации общества. При этом наблюдается обратная связь между заболеваемостью и социально-демографическими показателями: численностью населения, уровнем безработицы и численностью среднего медицинского персонала.

Рост заболеваемости РМЖ в Приморском крае оказался значимо связан с увеличением доходов при сокращении численности населения, что согласуется с данными Росстата. С 2000 г. на 45% сократился уровень безработицы и увеличилась численность работников-совместителей. Всё это говорит о росте занятости населения, что

приводит к нехватке времени на полноценный отдых и заботу о своём здоровье. Сократилось количество людей, прошедших диспансеризацию (в 2019 г. оно составило 84% от плана). В структуре населения возросла доля лиц старше трудоспособного возраста, риск заболеть РМЖ у которых выше.

Для построения прогнозной модели с учётом выявленных взаимосвязей в данной работе использовали регрессионный анализ, методы которого широко применяются в биомедицинских исследованиях [18, 20]. Отличием предложенной модели является то, что она разработана на основе выделенных главных компонент. Полученные значения оценок качества говорят об адекватности регрессионной модели. Наибольшее влияние на заболеваемость РМЖ, согласно параметрам модели, оказывают показатели, характеризующие социально-экономические условия жизни населения края. Похожие выводы были сделаны представителями ВОЗ при анализе ситуации по заболеваемости раком в мире. Согласно исследованиям уровень дохода и образ жизни человека являются внешними факторами риска, достоверные знания о которых позволят снизить заболеваемость РМЖ [1].

Настоящее исследование отличается от других [10, 11, 13–17, 19–21] набором доступных показателей, рассматриваемым временным интервалом, территориальной принадлежностью данных и использованием метода главных компонент при разработке прогнозной модели. Построенная модель позволяет прогнозировать заболеваемость РМЖ на территории Приморского края с учётом статистических взаимосвязей с факторами среды обитания и обра-

зом жизни человека. При этом динамика заболеваемости характеризуется линейным уравнением с высокой долей объяснённой дисперсии, что наблюдается в исследованиях других авторов. В работе А.В. Климушкина и соавт. при прогнозировании динамического ряда заболеваемости РМЖ в Оренбургской области получена модель с коэффициентом детерминации, равным 0,85 [25]. С помощью методов машинного обучения разработаны модели заболеваемости раком на европейском континенте, рассчитанные коэффициенты детерминации для которых описывают 95–99% дисперсии резульативного признака при низких значениях среднеквадратических ошибок [20].

Необходимость научных исследований в данной области диктуется социально-ориентированной политикой государства, реализуемой в настоящее время. Представляет интерес перекрёстная проверка качества предложенной модели по данным для других регионов России.

Ограничения исследования. Материалы исследования ограничены результатами статистического анализа 17 показателей КЖ населения Приморского края за 1994–2020 гг. и применением разработанной регрессионной модели на региональном уровне.

Выводы

Результаты данного исследования позволили выявить взаимосвязь заболеваемости РМЖ с факторами риска и разработать прогнозную модель, что может быть полезным при планировании профилактических мероприятий по улучшению КЖ и снижению заболеваемости на региональном уровне.

ЛИТЕРАТУРА

(п.п. 1, 2, 10–13, 15–17, 19–21 см. References)

3. Федеральная служба государственной статистики. Доступно: <https://rosstat.gov.ru>
4. Росстат. Российский статистический ежегодник – 2020. М.: 2020.
5. Нурманова А., Султанова З.И., Аннаоразов Ы.А. Факторы и их роль в заболеваемости, смертности, выживаемости при раке молочной железы. *Вестник Казахского национального медицинского университета*. 2018; (1): 112–4. <https://elibrary.ru/upzkou>
6. Одинцова И.Н., Писарева Л.Ф., Ананина О.А., Панферова Е.В. Заболеваемость раком молочной железы женского населения Сибирского федерального округа. *Здравоохранение Российской Федерации*. 2017; 61(3): 123–8. <https://doi.org/10.18821/0044-197X-2017-61-3-123-128> <https://elibrary.ru/yslavp>
7. Синкина Т.В., Петрова В.Д., Лазарев А.Ф. Современные представления о факторах риска рака молочной железы. *Российский биотерапевтический журнал*. 2009; 8(1): 88–94. <https://elibrary.ru/kxspz>
8. Французова И.С. Анализ факторов риска развития рака молочной железы. *Международный научно-исследовательский журнал*. 2019; (3): 68–74. <https://doi.org/10.23670/IRJ.2019.81.3.011> <https://elibrary.ru/yzpjm1>
9. Ширлина Н.Г., Стасенко В.Л., Ширинский В.А., Обухова Т.М. Распространенность факторов риска развития рака молочной железы в различных возрастных группах женского населения Омской области. *Эпидемиология и вакцинопрофилактика*. 2017; 16(6): 40–6. <https://elibrary.ru/zxrc1v>
14. Ширлина Н.Г. Прогнозирование риска развития рака молочной железы с помощью компьютерной программы. *Вестник СурГУ. Медицина*. 2016; (4): 63–6. <https://elibrary.ru/vlhyuz>
18. Кондратьев М.А. Методы прогнозирования и модели распространения заболеваний. *Компьютерные исследования и моделирование*. 2013; 5(5): 863–82. <https://elibrary.ru/rvbmux>
22. GitHub. Шитиков В.К., Мастицкий С.Э. Классификация, регрессия и другие алгоритмы DataMining с использованием R. Доступно: <https://github.com/ranalytics/data-mining>
23. Кабаков Р.И. *R в действии. Анализ и визуализация данных в программе R*. Пер. с англ. М.: ДМК Пресс; 2014.
24. Егшин В.Л., Саввина Н.В., Гржибовский А.М. Анализ главных компонент и факторный анализ в программной среде R. *West Kazakhstan Medical Journal*. 2020; 62(1): 6–14. <https://elibrary.ru/gfcmwd>
25. Климушкин А.В., Борщук Е.Л., Берун Д.Н., Берун Т.В., Куланова А.М. Прогноз заболеваемости злокачественными новообразованиями в Оренбургской области. *Современные проблемы науки и образования*. 2021; (2). <https://doi.org/10.17513/spno.30582> <https://elibrary.ru/ppqvhv>

REFERENCES

1. WHO. Cancer; 2022. Available at: <https://who.int/ru/news-room/fact-sheets/detail/cancer>
2. Siegel R.L., Miller K.D., Jemal A. Cancer statistics, 2020. *CA Cancer J. Clin.* 2020; 7(1): 7–30. <https://doi.org/10.3322/caac.21590>
3. Federal State Statistics Service. Available at: <https://rosstat.gov.ru>
4. Rosstat. Russian Statistical Yearbook – 2020. Moscow; 2020. (in Russian)
5. Nurmanova A., Sultanova Z.I., Annaorazov Y.A. Factors and their role in morbidity, mortality, survival in mammary cancer. *Vestnik Kazakhskogo natsional'no meditsinskogo universiteta*. 2018; (1): 112–4. <https://elibrary.ru/upzkou> (in Russian)
6. Odintsova I.N., Pisareva L.F., Ananina O.A., Panferova E.V. The morbidity of mammary gland cancer of female population of the Siberian federal district. *Zdravookhranenie Rossiyskoy*

- Federatsii*. 2017; 61(3): 123–8. <https://doi.org/10.18821/0044-197X-2017-61-3-123-128> <https://elibrary.ru/yslavp> (in Russian)
7. Sinkina T.V., Petrova V.D., Lazarev A.F. Morden view of breast cancer risk factors. *Rossiyskiy bioterapevicheskiy zhurnal*. 2009; 8(1): 88–94. <https://elibrary.ru/kxspz> (in Russian)
 8. Frantsuzova I.S. Analysis of risk factors of breast cancer development. *Mezhdunarodnyy nauchno-issledovatel'skiy zhurnal*. 2019; (3): 68–74. <https://doi.org/10.23670/IRJ.2019.81.3.011> <https://elibrary.ru/yzpjm1> (in Russian)
 9. Shirlina N.G., Stasenko V.L., Shirinskiy V.A., Obukhova T.M. The prevalence of risk factors of breast cancer in different age groups of the female population of the Omsk region. *Epidemiologiya i vaksinoprofilaktika*. 2017; 16(6): 40–6. <https://elibrary.ru/zxrc1v> (in Russian)
 10. Breccelmans C.T. Risk factors and risk reduction of breast and ovarian cancer. *Curr. Opin. Obstet. Gynecol.* 2003; 15(1): 63–8. <https://doi.org/10.1097/00001703-200302000-00010>
 11. Freedman D.M., Dosemeci M., McGlynn K. Sunlight and mortality from breast, ovarian, colon, prostate, and non-melanoma skin cancer: a composite death certificate based case-control study. *Occup. Environ. Med.* 2002; 59(4): 257–62. <https://doi.org/10.1136/oem.59.4.257>
 12. Kolpak E. Causes of morbidity and mortality from breast cancer. *Eurasian J. Anal. Chem.* 2017; 12(5): 779–91. <https://doi.org/10.12973/ejac.2017.00210a>
 13. Tyrer J., Duffy S.W., Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. *Stat. Med.* 2004; 23(7): 1111–30. <https://doi.org/10.1002/sim.1668>
 14. Shirlina N.G. Computer-aided breast cancer risk assessment. *Vestnik SurGU. Meditsina*. 2016; (4): 63–6. <https://elibrary.ru/vlhyyz> (in Russian)
 15. Xu J., Lin Y., Yang M., Zhang L. Statistics and pitfalls of trend analysis in cancer research: a review focused on statistical packages. *J. Cancer*. 2020; 11(10): 2957–61. <https://doi.org/10.7150/jca.43521>
 16. Lee S.J., Zelen A.M. Stochastic model for predicting the mortality of breast cancer. *J. Natl Cancer Inst. Monogr.* 2006; (36): 79–86. <https://doi.org/10.1093/jncimoographs/lgj011>
 17. Elwood M., Tawfiq E., TinTin S., Marshall R.J., Phung T.M., Campbell I., et al. Development and validation of a new predictive model for breast cancer survival in New Zealand and comparison to the Nottingham prognostic index. *BMC Cancer*. 2018; 18(1): 897. <https://doi.org/10.1186/s12885-018-4791-x>
 18. Kondrat'ev M.A. Forecasting methods and models of disease spread. *Komp'yuternye issledovaniya i modelirovanie*. 2013; 5(5): 863–82. <https://elibrary.ru/rvbmux> (in Russian)
 19. Van Maaren M.C., Kneepkens R.F., Verbaan J., Huijgens P.C., Lemmens V.E.P.P., Verhoeven R.H.A., et al. A conditional model predicting the 10-year annual extra mortality risk compared to the general population: A large population-based study in Dutch breast cancer patients. *PLoS One*. 2019; 14(1): e0210887. <https://doi.org/10.1371/journal.pone.0210887>
 20. Sekeroglu B., Tuncal K. Prediction of cancer incidence rates for the European continent using machine learning models. *Health Informatics J.* 2021; 27(1): 1460458220983878. <https://doi.org/10.1177/1460458220983878>
 21. Jafarabadi M.A., Iraj Z., Dolatkah R., Koshki T.J. Modeling the factors associated with mortality in patients with breast cancer: a machine learning approach. *Research Square*. 2020. Preprint. <https://doi.org/10.21203/rs.3.rs-57685/v1>
 22. GitHub. Shitikov V.K., Mastitskiy S.E. *Classification, Regression and Other DataMining Algorithms Using R [Klassifikatsiya, regressiya i drugie algoritmy DataMining s ispol'zovaniem R]*. Available at: <https://github.com/ranalytics/data-mining> (in Russian)
 23. Kabakov R.I. *R in Action. Data Analysis and Graphics with R [R v deystvii. Analiz i vizualizatsiya dannykh v programme R]*. Moscow: DMK Press; 2014. (in Russian)
 24. Egoshin V.L., Savvina N.V., Grzhibovskiy A.M. Principal components analysis and factor analysis in R. *West Kazakhstan Medical Journal*. 2020; 62(1): 6–14. <https://elibrary.ru/gfcmwd> (in Russian)
 25. Klimushkin A.V., Borshchuk E.L., Begun D.N., Begun T.V., Kulanova A.M. Forecast cancer incidence in Orenburg region. *Sovremennye problemy nauki i obrazovaniya*. 2021; (2). <https://doi.org/10.17513/spno.30582> <https://elibrary.ru/ppqwhw> (in Russian)

Информация об авторах

Ермолицкая Марина Захаровна — канд. биол. наук, ст. науч. сотр. лаб. информационно-аналитических и управляющих систем и технологий Института автоматизации и процессов управления ДВО РАН, 690041, Владивосток, Россия. E-mail: ermz@mail.ru

Кикун Павел Федорович — доктор мед. наук, канд. тех. наук, профессор, Дальневосточный федеральный университет, 690922, Владивосток, Россия.

Абакумов Александр Иванович — доктор ф.-м. наук, профессор, зав. лаб. математического моделирования биофизических процессов Института автоматизации и процессов управления ДВО РАН, 690041, Владивосток, Россия. E-mail: abakumov@dvo.ru

Information about the authors

Marina Z. Ermolitskaya — Ph.D. biol. Sciences, Art. scientific co-workers lab. Information-analytical and control systems and technologies, Institute of Automation and Control Processes, Far Eastern Branch of the Russian Academy of Sciences, Vladivostok, 690041, Russian Federation. E-mail: ermz@mail.ru <https://orcid.org/0000-0003-2588-102X>

Pavel F. Kiku — MD, PhD, DSci., Prof., Far Eastern Federal University, Vladivostok, 690922, Russian Federation. <https://orcid.org/0000-0003-3536-8617>

Alexander I. Abakumov — MD, PhD, DSci., Prof., Head of the Lab. of mathematical modelling of biophysical processes, Institute of Automation and Control Processes, Far Eastern Branch of the Russian Academy of Sciences, Vladivostok, 690041, Russian Federation. E-mail: abakumov@dvo.ru <https://orcid.org/0000-0003-2235-9025>