

УДК 519.237:519.722:616-006
DOI

КОМПЛЕКСНОЕ ИССЛЕДОВАНИЕ ПРОБЛЕМЫ ЛУЧЕВОЙ ДИАГНОСТИКИ РАКА МОЛОЧНОЙ ЖЕЛЕЗЫ С ПРИМЕНЕНИЕМ СОВРЕМЕННОЙ ТЕХНОЛОГИИ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ И ВЫЧИСЛИТЕЛЬНОГО ЭКСПЕРИМЕНТА

^{1,2}Тырсин А.Н., ³Попова А.Ю.

¹ФГАОУ ВО «Уральский федеральный университет имени первого Президента России Б.Н. Ельцина», Екатеринбург, e-mail: at2001@yandex.ru;

²Научно-инженерный центр «Надежность и ресурс больших систем и машин» Уральского отделения Российской академии наук, Екатеринбург, e-mail: at2001@yandex.ru;

³ГАОУЗ «Свердловский областной онкологический диспансер», Екатеринбург

На сегодняшний день классификация рака молочной железы основана на пяти биологических подтипах опухоли, которые различаются по прогнозу, агрессивности и чувствительности к различным видам терапии. Биотип опухоли определяют с помощью иммуногистохимического анализа. Актуальной проблемой является переход к неинвазивному установлению биотипа рака молочной железы. К основным диагностическим методам визуализации относят рентгеновскую проекционную маммографию и ультразвуковое исследование. Цель статьи – комплексное исследование данной проблемы лучевой диагностики рака молочной железы с помощью современной технологии математического моделирования и вычислительного эксперимента. Исследование выполнено на основе рентгенологических и ультразвуковых характеристик опухолей. Рассмотрены два метода исследования – классификация данных по биотипам с помощью линейного дискриминантного анализа и проведение векторного энтропийного анализа показателей для пациенток из каждого биотипа. Исследование проводилось на выборке из 365 пациенток. Дискриминантный анализ не позволил корректно классифицировать данные по соответствующим биотипам главным образом по причине корреляций между используемыми рентгенологическими и ультразвуковыми показателями. Векторное энтропийное моделирование показало, что взаимосвязи между рентгенологическими и ультразвуковыми характеристиками у биотипов рака молочной железы значительно отличаются друг от друга. Были проанализированы пять различных энтропийных показателей. Установлено, что эти энтропийные показатели между всеми биотипами попарно различимы. Полученные результаты комплексного исследования позволяют рассчитывать на разработку количественных критериев неинвазивной диагностики биотипов рака молочной железы.

Ключевые слова: модель, энтропия, корреляция, рак молочной железы, биологические подтипы, радиомика

PROBLEM OF RADIATION DIAGNOSIS OF BREAST CANCER: COMPREHENSIVE STUDY BASED ON MODERN MATHEMATICAL MODELING TECHNOLOGY AND COMPUTATIONAL EXPERIMENT

^{1,2}Tyrsin A.N., ³Popova A.Yu.

¹Ural Federal University named after the first President of Russia B.N. Yeltsin, Yekaterinburg, e-mail: at2001@yandex.ru;

²Scientific-Engineering Center Reliability and Life of Large Systems and Machines, Ural Branch, Russian Academy of Science, Yekaterinburg, e-mail: at2001@yandex.ru;

³Sverdlovsk Regional Oncological Dispensary, Yekaterinburg

To date, the classification of breast cancer is based on five biological subtypes of the tumour, which differ in prognosis, aggressiveness and sensitivity to various types of therapy. The tumour biotype is determined by immunohistochemical analysis. An urgent problem is the transition to non-invasive breast cancer biotype detection. The main diagnostic imaging methods include X-ray projection mammography and ultrasound. The purpose of the article is a comprehensive study of this problem of radiation diagnostics of breast cancer using modern technology of mathematical modelling and computational experiment. The study is based on the radiological and ultrasound characteristics of tumours. Two research methods are considered: classification of data by biotypes using linear discriminant analysis and vector entropy analysis of indicators for patients from each biotype. The study was conducted on a sample of 365 patients. Discriminant analysis did not allow the data to be correctly classified into the corresponding biotypes, mainly due to correlations between the radiological and ultrasound parameters used. Vector entropy modelling has shown that the relationship between radiological and ultrasound characteristics in the biotypes of breast cancer differs significantly from each other. 5 different entropy indicators were analysed. It was established that these entropy indices are pairwise distinguishable between all biotypes. The results of a comprehensive study allow us to count on the development of quantitative criteria for the non-invasive diagnosis of breast cancer biotypes.

Keywords: model, entropy, correlation, breast cancer, biological subtypes, radiomics

Введение

Рак молочной железы (РМЖ) – одно из самых распространенных онкологических заболеваний у женщин. Согласно данным Всемирной организации здравоохранения, на него приходится каждый шестой случай заболевания раком среди женщин. В 2022 г. в мире было выявлено 2,3 млн новых случаев. В России ежегодно РМЖ обнаруживают у 70 тыс. женщин и более 20 тыс. умирают. Злокачественные новообразования молочной железы остаются огромной медико-социальной проблемой, нанося колоссальный ущерб на государственном, региональном и индивидуальном уровнях [1].

На сегодняшний день классификация РМЖ основана на пяти биологических подтипах (биотипов) опухоли, которые различаются по прогнозу, агрессивности и чувствительности к различным видам лекарственной терапии: 1 – люминальный А; 2 – люминальный В; 3 – люминальный В HER2+; 4 – нелюминальный HER2+; 5 – базальноподобный TNBC [2]. Своевременное установление подтипа РМЖ позволяет подобрать индивидуальную схему лечения заболевания и улучшает выживаемость и прогноз [3]. В настоящее время биотип опухоли определяют с помощью иммуногистохимического анализа, относящегося к инвазивным методам исследования [4]. Очевидно, что актуальной проблемой является переход к неинвазивному установлению биотипа РМЖ, основанному на использовании адекватных математических моделей, учитывающих основные особенности данных. Сложность решения этой проблемы лучевой диагностики рака молочной железы требует комплексного исследования с применением современной технологии математического моделирования и вычислительного эксперимента.

К основным диагностическим методам визуализации относят рентгеновскую проекционную маммографию и ультразвуковое исследование. В ряде публикаций [2–6] отмечается определенный прогресс в использовании данных неинвазивных инструментов диагностики и оценки биологии опухоли. Одним из развивающихся направлений в неинвазивной диагностике является использование энтропийного анализа медицинских изображений. Публикуется большое количество работ, например [7–9], можно отметить определенный успех на стадии обнаружения отдельных заболеваний, включая и онкологию.

Однако пока отсутствуют обоснованные количественные критерии установления биотипа РМЖ на ранней стадии заболевания. Это можно объяснить тем, что данная проблема является более сложной по сравнению с обнаружением по изображениям РМЖ. По нашему мнению, ее эффективное решение требует системного подхода к математическому моделированию [10–12].

Цель исследования – комплексное исследование проблемы лучевой диагностики рака молочной железы на основе рентгенологических и ультразвуковых характеристик опухолей с применением современной технологии математического моделирования и вычислительного эксперимента.

Материалы и методы исследования

Из пациентов, проходивших обследование в Свердловском областном онкологическом диспансере с 2020 г., была сформирована выборка результатов рентгенологического и ультразвукового обследований для 365 пациенток возраста от 31 до 86 лет. Диагностические показатели описаны во многих публикациях [2; 13; 14, с. 49–79]. Они приведены в табл. 1. Для определения биотипов РМЖ было проведено комплексное гистологическое и иммуногистохимическое исследование всех пациенток. В результате в выборке оказались больные из всех пяти биотипов, в которых имеется 107, 177, 27, 21, 33 наблюдения соответственно.

Исследование включает две задачи. Во-первых, исследуем возможность по векторам рентгенологических $\mathbf{R} = (R_1, \dots, R_6)^T$ и ультразвуковых $\mathbf{U} = (U_1, \dots, U_6)^T$ показателей корректно классифицировать больных в соответствии с установленными биотипами. Многие авторы отмечают влияние возраста T на особенности заболевания, поэтому его также будем учитывать в анализе.

Во-вторых, попробуем найти существенные различия между биотипами на основе анализа рентгенологических и ультразвуковых показателей пяти кластеров (групп) пациенток.

Одним из наиболее распространенных методов классификации данных по обучающей выборке является линейный дискриминантный анализ [15, с. 10–17]. Он основан на возможности разделения совокупности данных на L кластеров с помощью гиперплоскостей, на основе обучающей выборки данных. Здесь объединяем все показатели в одно множество (вектор):

$$\mathbf{X} = \mathbf{R} \cup \mathbf{U} \cup T = (R_1, \dots, R_6, U_1, \dots, U_6, T)^T = (X_1, \dots, X_{13})^T.$$

Таблица 1

Система рентгенологических и ультразвуковых показателей

Обозначение	Показатели маммографии
R_1	Ведущий симптом по маммографии: 1 – микрокальцинаты в молочной железе; 2 – наличие узлов в молочной железе; 3 – узлы и микрокальцинаты в молочной железе; 4 – фокальная асимметрия молочной железы; 5 – узлы отдельно, микрокальцинаты отдельно
R_2	Контуры: 1 – четкие ровные контуры; 2 – четкие неровные контуры; 3 – нечеткие контуры; 4 – контуры с тяжистостью
R_3	Максимальный размер уплотнения (в мм): 1 – 0...5; 2 – 6...10; 3 – 11...15; 4 – 16...20
R_4	Плотность образования: 1 – гиперденсная; 2 – изоденсная; 3 – гиподенсная
R_5	Втяжение соска: 0 – нет; 1 – есть
R_6	Утолщение кожи: 0 – нет; 1 – есть
	Показатели магнитно-резонансной томографии
U_1	Эхогенность: 1 – гипо; 2 – гипер; 3 – смешанная
U_2	Дистальная акустическая тень: 1 – есть; 2 – нет
U_3	Контуры: 1 – четкие; 2 – нечеткие
U_4	Кровоток: 1 – периферический; 2 – интранодулярный; 3 – смешанный
U_5	Эластотип по Tsukuba
U_6	Микрокальцинаты: 1 – есть; 2 – нет

Рассмотрим его использование на примере двух кластеров [16, с. 509–516]. В этом случае имеем уравнение разделяющей гиперплоскости

$$f(\mathbf{x}) = \sum_{j=1}^M a_j x_j = C,$$

где $f(\mathbf{x})$ – дискриминантная функция, M – число переменных (анализируемых показателей), C – константа дискриминации, $\mathbf{x} = (x_1, \dots, x_M)^T$. Если $f(\mathbf{x}^0) > C$, то наблюдение $\mathbf{x}^0$ относим к первому кластеру, если $f(\mathbf{x}^0) < C$, то ко второму.

Вектор коэффициентов $\mathbf{a} = \mathbf{S}_*^{-1}(\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2)$, где $\bar{\mathbf{X}}_k = (\bar{X}_{k,1}, \dots, \bar{X}_{k,M})^T$ – вектор средних значений показателей k -го кластера; $\mathbf{S}_* = (\mathbf{X}_1^T \mathbf{X}_1 + \mathbf{X}_2^T \mathbf{X}_2) / (n_1 + n_2 - 2)$ – объединенная ковариационная матрица; $\mathbf{X}_k$ – матрица значений показателей k -го кластера; n_k – число наблюдений в k -м кластере; $C = 0,5(\bar{f}_1 + \bar{f}_2)$, где $\bar{f}_k = \frac{1}{n_k} \sum_{i=1}^{n_k} f(\mathbf{x}_{k,i})$; $k = 1, 2$.

Модели на основе дискриминантного анализа и других методов классификации не учитывают многомерную корреляцию анализируемых показателей. Для учета корреляций нужно рассматривать не отдельные наблюдения, а в целом биотипы РМЖ. Одним из подходов, учитывающих такие корреляции, является векторное энтропийное моделирование [17, с. 94–99]. Отметим несколько публикаций, в которых оно было применено для различных задач медицин-

ской диагностики [18–20]. Этот подход основан на векторном представлении дифференциальной энтропии, введенной в [21] как энтропия непрерывного распределения многомерной непрерывной случайной величины.

Дифференциальная энтропия (далее – энтропия) M -мерного случайного вектора $\mathbf{X} = (X_1, \dots, X_M)$ равна [17, с. 94]:

$$H(\mathbf{X}) = H_V(\mathbf{X}) - G_R(\mathbf{X}), \quad (1)$$

где

$$H_V(\mathbf{X}) = \sum_{j=1}^M \ln(\sigma_j) + \sum_{j=1}^M \kappa_j - \text{энтропийная «мера хаотичности»,}$$

$G_R(\mathbf{Y}) = -\frac{1}{2} \sum_{k=2}^M \ln(1 - R_{X_k/X_1, \dots, X_{k-1}}^2)$ – энтропийная «мера самоорганизации», σ_j – среднеквадратическое отклонение случайной величины Y_j , $\kappa_j = H(Y_j / \sigma_j)$ – энтропийный показатель типа закона распределения случайной величины Y_j , $R_{Y_k/Y_1, \dots, Y_{k-1}}^2$ – индексы детерминации регрессионных зависимостей.

Согласно (1): 1) если между компонентами случайного вектора $\mathbf{X}$ отсутствует строгая функциональная связь, но хотя бы одна компонента коррелирована X_j хотя бы с одной компонентой X_k ($j \neq k$), то $0 < G_R(\mathbf{X}) < \infty$; 2) если хотя бы одна компонента X_j строго функционально связана хотя бы с одной компонентой X_k ($j \neq k$), то $G_R(\mathbf{X}) = \infty$; 3) если все компоненты случайного вектора $\mathbf{X}$ взаимно независимы, то $G_R(\mathbf{X}) = 0$; 4) чем

теснее корреляционная связь между компонентами случайного вектора $\mathbf{X}$, тем больше значение $G_R(\mathbf{X})$. В частном случае, когда $\mathbf{X}$ – гауссов случайный вектор, имеем

$$H_V(\mathbf{X}) = \sum_{j=1}^M \ln(\sigma_j) + M \ln \sqrt{2\pi e},$$

$$G_R(\mathbf{Y}) = -\frac{1}{2} \ln |\mathbf{R}_X|, \quad (2)$$

где $\mathbf{R}_X$ – корреляционная матрица случайного вектора $\mathbf{X}$.

Энтропия (1) состоит из двух составляющих. Первое слагаемое $H_V(\mathbf{X})$ определяет предельную энтропию, соответствующую полной независимости показателей $\mathbf{X}$, и характеризует рассмотрение системы (РМЖ) как состоящего из частей (аддитивность). Второе слагаемое $G_R(\mathbf{X})$ отражает степень взаимосвязей между показателями, характеризуя свойства системы как целого (целостность). Поскольку $H_V(\mathbf{X})$ и $G_R(\mathbf{X})$ характеризуют разные стороны поведения заболевания как системы, то будем рассма-

тривать энтропию (1) в векторной форме в виде двух компонент – энтропийных мер хаотичности и самоорганизации:

$$h(\mathbf{X}) = (h_V, h_R) = (H_V(\mathbf{X}); G_R(\mathbf{X})). \quad (3)$$

Векторная энтропийная модель (1)–(3) учитывает все основные закономерности сложных систем [17, с. 34–43].

Рентгенологические и ультразвуковые показатели представлены в виде балльных значений. Поэтому при вычислении дифференциальных энтропий $H_V(X)$ используем методику, приведенную в [22], а при вычислении парных корреляций между показателями считаем, что для каждого балльного значения соответствующие фактические значения показателя имели некоторое гауссово распределение [23, с. 54–59].

Группы рентгенологических и ультразвуковых показателей можно интерпретировать как подсистемы в системе всех показателей. Поэтому для диагностирования биотипа РМЖ можно использовать энтропийную меру связи между этими группами [17, с. 97–98]:

$$G(\mathbf{R} \cap \mathbf{U}) = H(\mathbf{R}) + H(\mathbf{U}) - H(\mathbf{R} \cup \mathbf{U}) = -\frac{1}{2} \ln \frac{1 - d_e(\mathbf{R} \cup \mathbf{U})}{(1 - d_e(\mathbf{R}))(1 - d_e(\mathbf{U}))}, \quad (4)$$

где $d_e(\mathbf{R}) = 1 - \prod_{j=1}^6 (1 - R_{R_k/R_1, \dots, R_{k-1}}^2)$, $d_e(\mathbf{U}) = 1 - \prod_{j=1}^6 (1 - R_{U_k/U_1, \dots, U_{k-1}}^2)$ – коэффициенты тесноты корреляционной связи между компонентами случайных векторов $\mathbf{R}$ и $\mathbf{U}$.

Если $\mathbf{R}$ и $\mathbf{U}$ – гауссовы случайные векторы, имеем

$$G(\mathbf{R} \cap \mathbf{U}) = -\frac{1}{2} \ln \frac{|\mathbf{R}_{\mathbf{R} \cup \mathbf{U}}|}{|\mathbf{R}_R| \cdot |\mathbf{R}_U|}. \quad (5)$$

Согласно (4): 1) если между компонентами случайных векторов $\mathbf{R}$ и $\mathbf{U}$ отсутствует строгая функциональная связь, но хотя бы одна компонента коррелирована R , хотя бы с одной компонентой U_k , то $0 < G(\mathbf{R} \cap \mathbf{U}) < \infty$; 2) если хотя бы одна компонента R_i строго функционально связана хотя бы с одной компонентой U_k , то $G(\mathbf{R} \cap \mathbf{U}) = \infty$; 3) если все компоненты случайного вектора $\mathbf{R}$ взаимно независимы с компонентами случайного вектора $\mathbf{U}$, то $G(\mathbf{R} \cap \mathbf{U}) = 0$; 4) чем теснее корреляционная связь между компонентами случайных векторов $\mathbf{R}$ и $\mathbf{U}$, тем больше значение $G(\mathbf{R} \cap \mathbf{U})$.

Результаты исследования и их обсуждение

Выполним классификацию больных по пяти биотипам на основе линейного дискриминантного анализа. Для сопоставимости данных нормализуем показатель T (возраст пациентки) – приведем выборку

к минимуму, равному 0, и единичной дисперсии. Решение состоит в нахождении такой совокупности показателей, которая позволила бы (если это возможно) статистически достоверно распознать различия в этих группах. В табл. 2 представлены результаты дискриминантного анализа пяти биологических подтипов больных.

Параметр *Wilks' lambda* (λ , лямбда Уилкса, статистика Уилкса) принимает значения от 0 до 1. Чем меньше λ , тем лучше разделяются данные на пять кластеров. Значения параметра в целом по показателям $\lambda = 0,509$, что свидетельствует о неудовлетворительном результате классификации данных. В первом столбце (*Wilks' lambda*) табл. 3 приведены значения статистики Уилкса с исключенной соответствующей переменной. Во втором столбце (*Partial Lambda*) приведены значения частной лямбды Уилкса. Данная статистика описывает одиночный вклад соответствующей переменной в дискриминацию между совокупностями. Чем она меньше, тем больше вклад показателя в общую дискриминацию, видим, что ни одна из переменных не позволяет разделить совокупность многомерных данных на пять групп.

Таблица 2

Результаты дискриминантного анализа, $\lambda = 0,509$; $F(40, 1332) = 6,489$; $p < 0,000$

Показатель	Wilks' Lambda	Partial Lambda	F-remove (4, 351)	p-level	Toler.
R_2	0,565	0,901	9,677	0,000	0,734
R_3	0,563	0,904	9,312	0,000	0,925
R_4	0,538	0,946	4,984	0,001	0,944
R_5	0,529	0,962	3,444	0,009	0,768
R_6	0,534	0,954	4,269	0,002	0,771
U_1	0,530	0,960	3,633	0,006	0,919
U_2	0,531	0,958	3,818	0,005	0,741
U_3	0,534	0,953	4,319	0,002	0,873
U_6	0,532	0,958	3,853	0,004	0,926
T	0,565	0,901	9,669	0,000	0,911

Источник: составлено авторами на основе полученных данных в ходе исследования.

Таблица 3

Результаты классификации данных на пять биотипов

Group	Percent Correct	G 1:1 $p = 0,293$	G 2:2 $p = 0,485$	G 3:3 $p = 0,074$	G 4:4 $p = 0,058$	G 5:5 $p = 0,090$
G 1:1	36,45	39	61	0	4	3
G 2:2	85,31	12	151	6	0	8
G 3:3	55,56	0	12	15	0	0
G 4:4	4,76	6	14	0	1	0
G 5:5	54,55	0	14	0	1	18
Total	61,37	57	252	21	6	29

Источник: составлено авторами на основе полученных данных в ходе исследования.

В следующих двух столбцах приведены F -статистики (F -remove) дисперсионного анализа и соответствующие p -уровни, характеризующие вклад каждой из переменных в различие двух групп. В соответствии с результатами наиболее значимыми являются показатели R_2 , R_3 , T . Значение толерантности ($Toler$) показателя вычисляется как $1 - R^2$, где R^2 – коэффициент детерминации линейной регрессии этого признака X_i на оставшуюся совокупность статистически значимых признаков, использованную в анализе.

В табл. 3 приведены результаты классификации данных.

Из табл. 3 видим, что только 61,4% пациенток были правильно классифицированы. А для биотипа 4 правильно классифицирована только одна пациентка из 21. Таким образом можно говорить о неудовлетворительной классификации больных на биотипы РМЖ.

Энтропийный анализ проведем по всем популяциям пациенток для всех пяти биологических подтипов. Из-за линейной зависимости и постоянства значений для всех пациенток биотипа исключим по два показателя из каждой группы. В результате будем рассматривать четыре рентгенологических

$\mathbf{R} = (R_1, R_2, R_3, R_4)^T$ и четыре ультразвуковых $\mathbf{U} = (U_1, U_2, U_3, U_6)^T$ показателя. Для энтропийного анализа в рамках соотношений (1)–(5) достаточно для всех пяти биотипов определить корреляционные матрицы и вычислить среднеквадратические отклонения рентгенологических и ультразвуковых показателей. Это легко реализуется даже средствами MS Excel.

В табл. 4 приведены рассчитанные по формулам (2), (4), (5) энтропийные показатели для всех биотипов РМЖ. Видим, что во многих случаях показатели для разных биотипов существенно различаются. Поэтому сравним попарно показатели между биотипами. Будем считать, что значения показателя у двух биотипов существенно различимы, если они отличаются хотя бы в 1,9 раза.

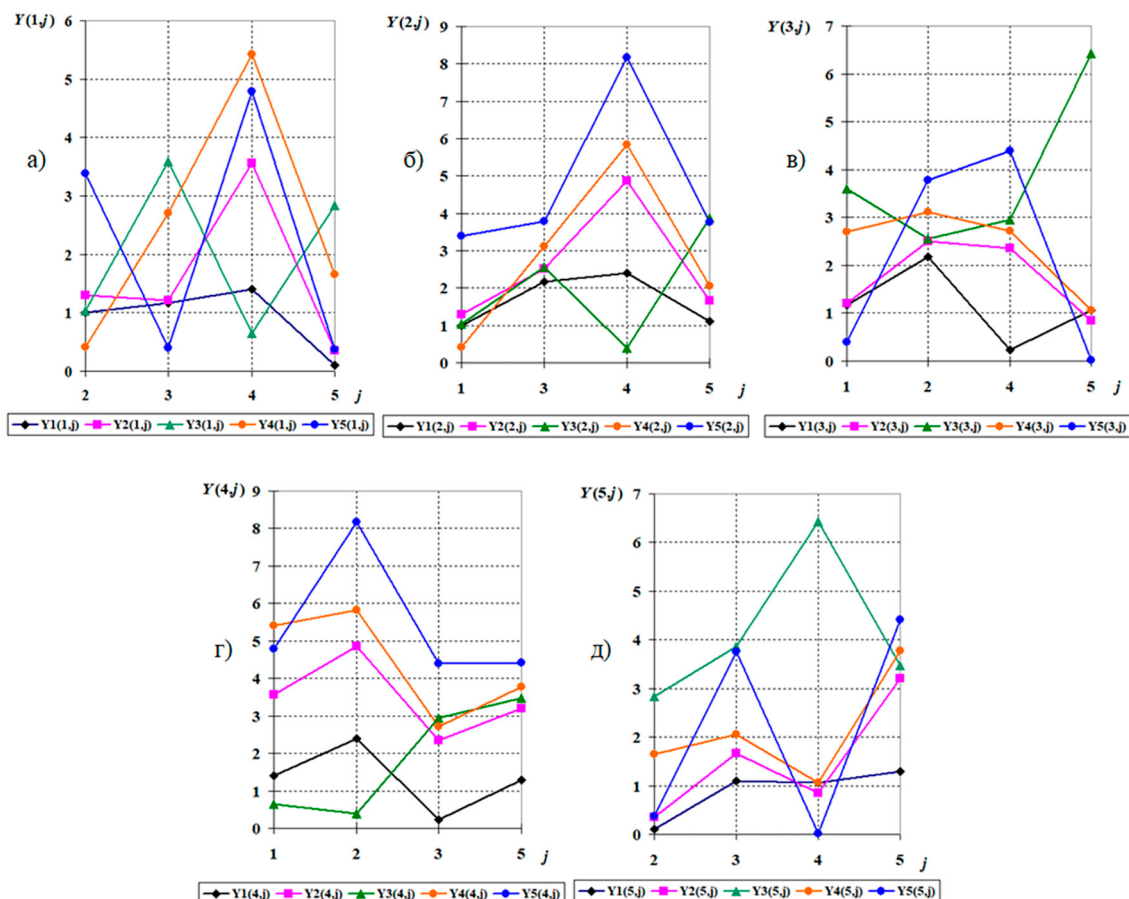
На рисунке для каждого из подтипов приведены значения всех пяти показателей $Y_{(k,j)}$, рассчитанных по табл. 4. Логарифмирование позволяет легко интерпретировать рисунки – если значение показателя не ниже 1, то это означает существенное различие между соответствующими подтипами. В табл. 5 дана сводка значимых попарных отношений энтропийных показателей из рисунка.

Таблица 4

Энтропийные показатели для всех биотипов

Подтип	$G_r(\mathbf{R} \cup \mathbf{U})$	$G_r(\mathbf{R})$	$G_r(\mathbf{U})$	$G(\mathbf{R} \cap \mathbf{U})$	$H_v(\mathbf{R})$	$H_v(\mathbf{U})$	$H_v(\mathbf{R}) - H_v(\mathbf{U})$
1	0,843	0,065	0,225	0,554	3,576	2,625	0,951
2	0,368	0,050	0,026	0,292	3,143	1,306	1,837
3	1,823	0,365	0,289	1,168	3,371	-6,062	9,434
4	8,218	2,065	4,802	1,351	3,936	2,499	1,436
5	1,061	0,186	0,286	0,590	2,769	2,613	0,156

Источник: составлено авторами на основе полученных данных в ходе исследования.

Значения всех пяти показателей $Y_i(k, j)$ для каждого из подтипов

$$\text{Обозначим } Y_1(k, j) = \log_a \left(\max_{k, j} \frac{G(\mathbf{R}_k \cap \mathbf{U}_k)}{G(\mathbf{R}_j \cap \mathbf{U}_j)} \right), \quad Y_2(k, j) = \log_a \left(\max_{k, j} \frac{G_r(\mathbf{R}_k \cup \mathbf{U}_k)}{G_r(\mathbf{R}_j \cup \mathbf{U}_j)} \right),$$

$$Y_3(k, j) = \log_a \left(\max_{k, j} \frac{H_v(\mathbf{R}_k) - H_v(\mathbf{U}_k)}{H_v(\mathbf{R}_j) - H_v(\mathbf{U}_j)} \right), \quad Y_4(k, j) = \log_a \left(\max_{k, j} \frac{G_r(\mathbf{R}_k)}{G_r(\mathbf{R}_j)} \right),$$

$$Y_5(k, j) = \log_a \left(\max_{k, j} \frac{G_r(\mathbf{U}_k)}{G_r(\mathbf{U}_j)} \right), \quad a = 1,9.$$

Из табл. 5 видно, что от 2 до 5 раз энтропийные показатели между всеми биотипами различаются более чем в 2 раза.

Таблица 5

Сводка значимых попарных отношений
энтропийных показателей

$j \backslash k$	1	2	3	4	5
1	–	4	4	4	2
2	4	–	5	4	5
3	4	5	–	4	3
4	4	4	4	–	5
5	2	5	3	5	–

Источник: составлено авторами на основе полученных данных в ходе исследования.

Заключение

Проведенное комплексное исследование проблемы лучевой диагностики рака молочной железы с применением современной технологии математического моделирования и вычислительного эксперимента показало следующее. Неудачное использование линейной классификации данных свидетельствует о наличии корреляций между используемыми рентгенологическими и ультразвуковыми показателями. Векторное энтропийное моделирование показало, что множества рентгенологических и ультразвуковых показателей у пяти биотипов РМЖ имеют различия на системном уровне. Проанализированы пять энтропийных показателей. Оказалось, что не менее чем в двух случаях эти энтропийные показатели между всеми биотипами попарно различимы. Это позволяет в дальнейшем разработать количественные критерии неинвазивной диагностики биотипов РМЖ на начальной стадии заболевания. Результат может быть достигнут за счет совместного использования двух рассмотренных методов математического моделирования – линейного дискриминантного анализа и векторного энтропийного моделирования. Для повышения достоверности этих критериев в дальнейшем желательно увеличить обучающие выборки больных для 3, 4, 5-го биотипов, а также привести используемые в рамках исследования наборы данных.

Список литературы

1. Зайкова З.А., Кравченко Н.А. Рак молочной железы: статистика заболеваемости и смертности // Якутский медицинский журнал. 2023. № 3. С. 72–75. DOI: 10.25789/YMJ.2023.83.18.
2. Попова А.Ю., Гаждонова В.Е., Карташов М.В., Шевченко С.А., Белова О.С. Радиомика в лучевой диагностике биологических подтипов рака молочной железы (обзор литературы) // Уральский медицинский журнал. 2024. Т. 23. № 4. С. 41–56. DOI: 10.52420/umj.23.4.41.
3. Ginsburg O., Yip C., Brooks A., Cabanes A., Caleffi M., Yataco J., Gyawali B. Breast cancer early detection: A phased approach to implementation // Cancer. 2020. Vol. 126. Suppl. 10. P. 2379–2393. DOI: 10.1002/cnecr.32887.
4. Johnson K.S., Conant E.F., Soo M.S. Molecular subtypes of breast cancer: A review for breast radiologists // Journal of Breast Imaging. 2021. Vol. 3, Is. 1. P. 12–24. DOI: 10.1093/jbi/wbaa110.
5. Shaikh S., Rasheed A. Predicting Molecular Subtypes of Breast Cancer with Mammography and Ultrasound Findings: Introduction of Sono-Mammometry Score // Radiology Research and Practice. Vol. 2021. Art. ID 6691958. 12 p. DOI: 10.1155/2021/6691958.
6. Sezgin G., Apaydin M., Ettit D., Atahan M.K. Tumour size estimation of the breast cancer molecular subtypes using imaging techniques // Medicine and Pharmacy Reports. 2020. Vol. 93, Is. 3. P. 253–259. DOI: 10.15386/mpr-1476.
7. Ming-Yih Lee, Chi-Shih Yang. Entropy-based feature extraction and decision tree induction for breast cancer diagnosis with standardized thermograph images // Computer Methods and Programs in Biomedicine. 2010. Vol. 100. № 3. P. 269–282. DOI: 10.1016/j.cmpb.2010.04.014.
8. Razmjoooy N., Estrela V.V., Loschi H.J. Entropy-Based Breast Cancer Detection in Digital Mammograms Using World Cup Optimization Algorithm // International Journal of Swarm Intelligence Research. 2020. Vol. 11, Is. 3. P. 1–18. DOI: 10.4018/ijisir.2020070101.
9. Palczynski K., Ledzinski D., Andrysiak T. Entropy Measurements for Leukocytes' Surrounding Informativeness Evaluation for Acute Lymphoblastic Leukemia Classification // Entropy. 2022. Vol. 24. P. 1560. DOI: 10.3390/e24111560.
10. Карякина О.Е., Добродева Л.К., Мартынова Н.А., Красильников С.В., Карякина Т.И. Применение математических моделей в клинической практике // Экология человека. 2012. Т. 19. № 7. С. 55–64. DOI: 10.17816/humeco17447.
11. Бобков С.П., Астраханцева И.А., Галиаскаров Э.Г. Применение системного подхода при разработке математических моделей // Современные наукоемкие технологии. Региональное приложение. 2021. № 1 (65). С. 66–71. DOI: 10.6060/snt.20216501.0008.
12. Cassidy R., Singh N.S., Schiratti P.-R., Semwanga A., Binyaruka P., Sachingongu N., Chama-Chiliba Ch.M., Chalabi Z., Borghi J., Blanchet K. Mathematical modelling for health systems research: a systematic review of system dynamics and agent-based models // BMC Health Serv Res. 2019. Vol. 19. P. 845. DOI: 10.1186/s12913-019-4627-7.
13. Алиева Г.С., Корженкова Г.П., Колядина И.В. Комплексная лучевая диагностика раннего рака молочной железы (обзор литературы) // Современная онкология. 2019. Т. 21. № 3. С. 26–32. DOI: 10.26442/18151434.2019.3.190469.
14. Сенча А.Н., Бикеев Ю.В. Ультразвуковое исследование молочных желез. Атлас. М.: МЕДпресс-информ, 2021. 296 с. ISBN 978-5-00030-986-5.
15. Shinmura S. New Theory of Discriminant Analysis after R. Fisher. Springer. 2016. 220 p. DOI: 10.1007/978-981-10-2164-0.
16. Сошникова Л.А., Тамашевич В.Н., Уебе Г., Шефер М. Многомерный статистический анализ в экономике. М.: Финансы и статистика, 1999. 598 с. ISBN 5-238-00099-5.
17. Тырсин А.Н. Векторное энтропийное моделирование многомерных стохастических систем. М.: Наука, 2022. 231 с. [Электронный ресурс]. URL: https://elar.urfu.ru/bitstream/10995/119918/1/978-5-02-040905-7_2022.pdf (дата обращения: 28.05.2025). ISBN 978-5-02-040905-7.
18. Tyrsin A.N., Gevorgyan G.G. Population Health Assessment Based on Entropy Modeling of Multidimensional Stochastic Systems // Communications in Computer and Information Science. 2019. Vol. 794. P. 92–105. DOI: 10.1007/978-3-030-35400-8_7.
19. Яшин Д.А., Тырсин А.Н. Энтропийно-вероятностный анализ безопасности лекарственных средств на примере терапии липиднормализующими препаратами // Современные проблемы науки и образования. 2021. № 2. URL: <http://www.science-education.ru/ru/article/view?id=30614> (дата обращения: 30.05.2025). DOI: 10.17513/spno.30614.
20. Тырсин А.Н., Улезко Е.А., Остроушко Д.В., Свирская О.Я., Санковец Д.Н. Логистическая регрессия как модель оценки состояния недоношенных новорожденных с целью их перевода на неинвазивную респираторную поддержку // Современные наукоемкие технологии. 2022. № 3. С. 49–55. DOI: 10.17513/snt.39072.
21. Shannon C.E. A Mathematical Theory of Communication // The Bell System Technical Journal. 1948. Vol. 27. P. 379–423, 623–656. URL: <https://archive.org/details/bstj27-3-379/mode/2up> (дата обращения: 30.05.2025).
22. Тырсин А.Н. Энтропийное моделирование дискретных случайных векторов на примере группировок и балльных показателей // Современные наукоемкие технологии. 2021. № 1. С. 51–56. DOI: 10.17513/snt.38470.
23. Ксендзенко Л.С., Бойко Л.А., Шнишкин А.В. Элементы теории корреляции. Владивосток: ДВФУ, 2025. 71 с. ISBN 978-5-7444-5829-4.