



DOI: <https://doi.org/10.15688/jvolsu1.2016.6.12>

УДК 618.19:616-006+616-073

ББК 55.6-4

## ОБ ОПТИМИЗАЦИИ ЧИСЛА ДИАГНОСТИЧЕСКИХ ПРИЗНАКОВ ЗАБОЛЕВАНИЙ МОЛОЧНЫХ ЖЕЛЕЗ НА ОСНОВЕ ТЕРМОМЕТРИЧЕСКИХ ДАННЫХ<sup>1</sup>

**Елена Алексеевна Мазепа**

Кандидат физико-математических наук,  
доцент кафедры математического анализа и теории функций,  
Волгоградский государственный университет  
lmazepa@rambler.ru, matf@volsu.ru  
просп. Университетский, 100, 400062 г. Волгоград, Российская Федерация

**Хеди Мовладовна Сулейманова**

Аспирант кафедры математического анализа и теории функций,  
Волгоградский государственный университет  
hedi.medi@mail.ru, matf@volsu.ru  
просп. Университетский, 100, 400062 г. Волгоград, Российская Федерация

**Аннотация.** Работа посвящена, с одной стороны, исследованию корреляционной взаимосвязи признаков, полученных на основе медицинских термометрических данных, предназначенных для диагностики заболеваний молочных желез. Также выделены наиболее значимые признаки того или иного заболевания для экспресс-диагностики злокачественных новообразований молочной железы.

**Ключевые слова:** микроволновая радиотермометрия, консультативные интеллектуальные системы, экспресс-диагностика, злокачественные новообразования молочной железы, корреляционный анализ.

### Введение

Одним из направлений развития систем искусственного интеллекта является разработка экспертных систем медицинской диагностики. Их применение помогает врачу повысить качество своей работы. Задачей таких систем является не только определение, но и консультационная помощь в определении заболеваний (одного или нескольких), которые, возможно, имеет пациент, на основе данных о его наблюдениях. При создании консультационных интеллектуальных систем разработчики опираются на целый ряд качественных диагностических признаков, указанных врачами. Для их математической интерпретации могут быть использованы различные функциональные зависимости между исходными термометрическими данными. В связи с этим число возможных количественных диагностических признаков, которые могут быть применены в системах экспресс-диагностики, возрастает до нескольких сотен и даже тысяч. Каждый из полученных признаков оценивался с точки зрения его информативности, которая отражает степень диагностической способности этого признака. Однако многие из полученных ко-

личественных признаков оказываются взаимозависимыми и взаимозаменяемыми. Уменьшение числа диагностических признаков с учетом их независимости и максимальной информативности является одним из направлений повышения качества работы консультационных интеллектуальных систем.

**Цель исследования** – оптимизация количества диагностических признаков с учетом их некоррелированности и с сохранением их максимальной информативности.

Для достижения поставленной цели будут решены следующие задачи: определение степени корреляционной зависимости диагностических признаков друг от друга; выделение групп взаимозависимых признаков; нахождение внутригрупповых признаков с максимальной информативностью.

### Материалы и методы исследования

В настоящее время одной из актуальных проблем в медицине является ранняя диагностика заболеваний молочной железы. Рак молочной железы (РМЖ) занимает лидирующие позиции среди множества заболеваний у женщин, поскольку несет наибольшую опасность для здоровья среди злокачественных опухолей.

На сегодняшний момент диагностика заболеваний молочных желез основывается на осмотре и расспросе больного. В специализированных клиниках и кабинетах применяются такие инструментальные исследования, как ультразвуковое сканирование, рентгенологические методы и др. Для того чтобы своевременно выявлять большинство опухолей, они должны иметь диаметр не более 5–7 мм. Так как средний размер впервые выявляемых опухолей значительно больше (1,34 см), частота выявления опухолей до 1 см в диаметре, по мировым данным, составляет 10–20 %. Распознавание заболеваний на ранней стадии может приблизиться к идеальному при рентгенологическом обследовании молочных желез каждые 6 месяцев, однако это влечет за собой комплекс новых проблем. Среди них возможность индуцирования рака вследствие частого облучения ткани молочной железы и значительное увеличение стоимости профилактических программ [5]. Кроме того, применение этих методов требует подготовки высококвалифицированных кадров, наличия дорогостоящего оборудования и не решает проблему ранней диагностики и динамического контроля над течением заболеваний.

Одним из направлений решения указанной проблемы вполне может стать создание экспертных комплексов, диагностирующих РМЖ, которые должны обладать следующими свойствами: 1) невысокая цена; 2) простота в применении (чтобы работу с данным комплексом мог бы осуществлять средний медицинский персонал); 3) безопасность как для пациента, так и для медицинского персонала; 4) надежность проводимых измерений, а также высокое качество диагностики.

Например, работа такого комплекса может быть основана на применении метода комбинированной радиотермометрии (это биофизический метод неинвазивного обследования, который базируется на измерении температуры тела на глубине нескольких сантиметров (РТМ) и нажном покрове (ИК)). Не секрет, что при возникновении заболевания какого-либо органа происходит изменение температурного фона определенного участка тела человека.

Данную технологию в маммологии и других областях медицины рекомендуют для скрининга, дифференциальной диагностики и оценки эффективности проводимого лечения [6].

Опишем подробнее существующую методику диагностики рака молочной железы по данным микроволновой радиотермометрии. Обследование пациентки проводится в горизонтальном положении, обнаженной по пояс, руки под головой. Оно начинается с измерения температур в опорных точках T1 и T2, расположенных следующим образом: первая в центре грудной клетки сразу под и между молочными железами, вторая непосредственно под мечевидным отростком. Далее измерения проводятся в 10 точках на каждой железе, включая аксиллярные области (см. рисунок).

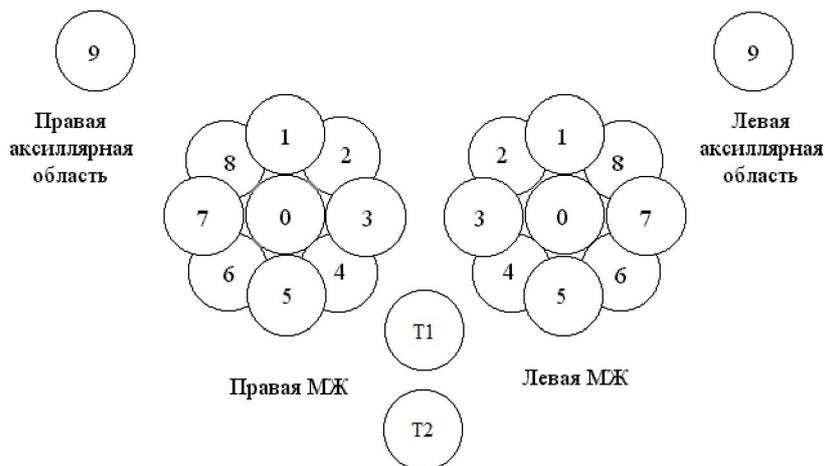


Схема обследования молочной железы

Все женщины были обследованы с использованием радиотермометрического компьютеризированного комплекса РТМ-01-РЭС (ООО «Фирма РЭС»), принцип действия которого основан на измерении собственного электромагнитного излучения тканей в микроволновом диапазоне (глубинная температура – РТМ) и инфракрасном диапазоне (температура кожи – ИК) [7].

В результате проведенных исследований была сформирована обучающая выборка, которая включает в себя информацию о 734 молочных железах пациенток. Было решено разделить ее на два контрольных класса: здоровые – 148 молочных желез и больные – 586 молочных желез.

При РТМ-обследовании РМЖ характеризуется повышением температуры в отдельных областях молочной железы из-за усиленного метаболизма (набор химических реакций, которые возникают в живом организме для поддержания жизни) раковых клеток.

### Качественные диагностические признаки и их количественные аналоги

В ходе исследований и анализа термометрических данных специалистами были выявлены следующие качественные признаки рака молочной железы. Все признаки были разбиты на пять основных групп [2].

1. Группа признаков, основанных на вычислении разности температур в различных точках или подобластях молочной железы.

1.1. Повышенное максимальное значение разности температур между одноименными точками правой и левой молочных желез. В качестве характеристик, описывающих этот эффект, могут быть использованы функции вида

$$|t_{i,np} - t_{i,л}|, \text{ или } t_{i,np} - t_{i,л},$$

где  $t_{i,np}$  и  $t_{i,л}$  – температуры в  $i$ -ных точках правой и левой молочных желез соответственно,  $i = 1, \dots, 8$ .

1.2. Повышенная разница температур сосков правой и левой молочных желез:

$$|t_{0,np} - t_{0,л}|.$$

1.3. Повышенное среднеквадратичное значение разностей температур между одноименными точками правой и левой молочных желез:

$$\sqrt{\sum_{i=0}^8 \frac{(t_{i,np} - t_{i,л})^2}{9}}.$$

1.4. Повышенное среднеквадратичное отклонение разностей температур между одноименными точками правой и левой молочных желез:

$$\sqrt{\sum_{i=0}^8 \frac{(\Delta t_i - \overline{t_{мжс}})^2}{9}},$$

где  $\Delta t_i = t_{i, np} - t_{i, л}$ ,  $\overline{t_{мжс}} = \sum_{i=0}^8 \frac{\Delta t_i}{9}$ .

1.5. Разница средних значений температур подобластей молочных желез, например:

$$\left| \frac{t_{0, PTM, np} + t_{i, PTM, np} + t_{i(\text{mod}8)+1, PTM, np}}{3} - \frac{t_{0, PTM, л} + t_{i, PTM, л} + t_{i(\text{mod}8)+1, PTM, л}}{3} \right|,$$

$$i = 0, \dots, 8.$$

1.6. Разница осцилляций подобластей молочных желез, например:

$$\left| (\max_{i=0,8} t_{i, PTM} - \min_{i=0,8} t_{i, PTM})_{np} - (\max_{i=0,8} t_{i, PTM} - \min_{i=0,8} t_{i, PTM})_{л} \right|.$$

1.7. Разница среднеквадратичных отклонений температур молочных желез:

$$\left| \left( \sum_{i=0}^8 \frac{(t_{i, PTM} - t_{PTM, cp})^2}{8} \right)_{np} - \left( \sum_{i=0}^8 \frac{(t_{i, PTM} - t_{PTM, cp})^2}{8} \right)_{л} \right|.$$

2. Группа признаков, характеризующих повышенный разброс температур в пораженной молочной железе.

2.1. Повышенное среднеквадратичное отклонение температур в одной из молочных желез:

$$\sqrt{\sum_{i=0}^8 \frac{(\Delta t_i - \overline{t_{cp}})^2}{9}}, \text{ где } \overline{t_{cp}} = \sum_{i=0}^8 \frac{t_i}{9}.$$

2.2. Повышенное значение осцилляции молочной железы или отдельных областей, например:

$$\max_{i=1,8} t_i - \min_{i=1,8} t_i.$$

2.3. Повышенный разброс температур между отдельными точками в пораженной молочной железе:

$$t_{cp} - t_i,$$

где  $t_{cp} = \sum_{i=1}^8 \frac{t_i}{8}$ ,  $t_i$  – температура в  $i$ -ной точке молочной железы.

2.4. Повышенные средние значения температур соседних точек вплоть до средней температуры молочной железы, то есть функции вида:

$$t_{cp} - \frac{t_i + t_{i(\text{mod}8)+1} + t_{i(\text{mod}8)+2}}{3},$$

$$t_{cp} - \frac{t_i + t_{i(\text{mod}8)+1} + t_{i(\text{mod}8)+2} + t_{i(\text{mod}8)+3}}{4} \text{ и т. д.}$$

3. Группа признаков, характеризующих повышенное значение температуры соска в пораженной молочной железе.

3.1. Аномальная разность температуры соска и средней температуры молочной железы:

$$t_0 - t_{cp}, \text{ где } t_{cp} = \sum_{i=1}^8 \frac{t_i}{8}.$$

3.2. Аномальная разность температуры соска и температур отдельных точек молочной железы:

$$t_0 - t_i, i = 0, \dots, 8.$$

3.3. Аномальная разность температуры соска и средней температуры различных областей молочной железы, например:

$$t_0 - \frac{t_i + t_{i(\text{mod}8)+1}}{2}, i = 1, \dots, 8 \text{ или}$$

$$t_0 - \frac{t_i + t_{i(\text{mod}8)+1} + t_{i(\text{mod}8)+2}}{3}, i = 1, \dots, 8 \text{ или}$$

$$t_0 - \frac{t_i + t_{i(\text{mod}8)+1} + t_{i(\text{mod}8)+2} + t_{i(\text{mod}8)+3}}{4}, i = 1, \dots, 8 \text{ и т. д.}$$

4. Группа признаков, характеризующих соотношение кожной и глубинной температур.

4.1. Аномальное значение разности между кожной и глубинной температурами точки пораженной молочной железы (внутренний градиент):

$$t_{i,PTM} - t_{i,ИК},$$

где  $t_{i,PTM}$  – глубинные и  $t_{i,ИК}$  – кожные температуры в  $i$ -ной точке молочной железы.

4.2. Аномальное значение разности кожной и глубинной средних температур соседних точек пораженной молочной железы:

$$\frac{(t_{i,PTM} - t_{i,ИК}) + (t_{i(\text{mod}8)+1,PTM} - t_{i(\text{mod}8)+1,ИК})}{2}, i = 1, \dots, 8.$$

5. Признаки, базирующиеся на различных параметрах физико-математических моделей поведения температурных полей.

При исследовании было рассмотрено всего порядка 800 числовых функций между различными исходными термометрическими данными.

### Ранговая корреляция Спирмена

В основе исследования связи между диагностическими признаками лежат статистические методы, основанные на вычислении выборочного коэффициента корреляции и на проверке гипотез о его значимости. При этом большинство параметрических критериев проверки статистических гипотез предполагает нормальность распределения выборочных данных. Для проверки

выборочных распределений на принадлежность к нормальному распределению могут быть использованы как приближенный метод проверки гипотезы о нормальном распределении, основанный на вычислении выборочных коэффициентов асимметрии и эксцесса, так и более точный метод проверки согласованности – критерий согласия Пирсона.

Ранее проводились аналогичные теоретические исследования, определяющие степень зависимости температурных изменений в молочной железе от диаметра железы, температуры в аксиллярной области, в контрольных точках на теле пациентки, а также от ряда анамнестических данных [1]. В указанной работе было показано, что большинство выборочных термометрических данных не удовлетворяет нормальному распределению. Поэтому для изучения взаимосвязи диагностических признаков друг с другом целесообразно применять ранговые методы исследования корреляционной взаимосвязи, в частности, основанные на вычислении коэффициента корреляции Спирмена.

В основе данного метода лежит вычисление коэффициентов ранговой корреляции Спирмена  $\rho_g$  по формуле [1; 3]:

$$\rho_g = 1 - \frac{6}{n^3 - n} \sum_{i=1}^n d_i^2, \quad (1)$$

где  $d_i = x_i - y_i$ , а  $x_i$  и  $y_i$  ( $i = 1, 2, \dots, n$ ) – ранги исходных объектов, записанные по исследуемым признакам.

Данный коэффициент корреляции позволяет определить, существует ли между двумя переменными зависимость, выражаемая монотонной функцией (то есть при росте одной переменной увеличивается и вторая, и наоборот).

Чтобы выяснить, насколько значение коэффициентов ранговой корреляции Спирмена  $\rho_g$  несет случайный или неслучайный характер, необходимо исследовать значимость полученных коэффициентов.

Для выбора гипотезы о значимости коэффициентов ранговой корреляции Спирмена использовались статистический критерий уровня  $\alpha = 0,05$  ( $p = 1 - \alpha$ ) и следующие значения критических точек распределения Стьюдента  $t_{кр}$  (табл. 1).

Таблица 1

Значения критических точек распределения Стьюдента,  $t_{кр}$ 

Молочные железы	$n$ (объем выборки)	$t_{кр}$
«Здоровые»	148	1,97
«Больные»	586	1,96

При этом вывод о значимости или незначимости корреляционной взаимосвязи между признаками делался в соответствии со следующим правилом: если  $|T_{набл}| < t_{кр}$ , то принимается гипотеза  $\rho_g = 0$ , то есть между признаками нет корреляционной зависимости; если  $|T_{набл}| > t_{кр}$ , то принимается гипотеза  $\rho_g \neq 0$ , что говорит о наличии корреляционной зависимости между признаками [1; 3; 4]. Заметим, что в данном правиле в качестве критерия используется статистика

$$T_{набл} = \frac{\rho_g}{\sqrt{1 - \rho_g^2}} \sqrt{n - 2}, \quad (2)$$

имеющая распределение Стьюдента с  $n - 2$  степенями свободы, где  $\rho_g$  – коэффициент ранговой корреляции Спирмена,  $n$  – число элементов выборки.

### Результаты и их обсуждения

Для каждой из выявленных групп признаков, описанных выше, по термометрическим данным были построены корреляционные матрицы. Они позволили в каждой группе выделить под-

группы коррелирующих друг с другом признаков. Для оценивания степени корреляционной взаимосвязи использовалась шкала Чеддока (табл. 2).

Таблица 2

Шкала Чеддока

Коэффициент	0,1–0,3	0,3–0,5	0,5–0,7	0,7–0,9	0,9–0,9
Характеристика зависимости	Слабая	Умеренная	Заметная	Высокая	Весьма высокая

В качестве порогового значения коэффициента корреляции было выбрано значение  $\rho_g \geq 0,7$ . Заметим, что при таком значении  $\rho_g$  выполнено неравенство  $|T_{набл}| > t_{кр}$ , то есть корреляционная взаимосвязь признаков является значимой. Ниже приведем таблицы 3–8, в которых представлены подгруппы признаков из первой группы с высокой корреляционной взаимосвязью, а также корреляционные матрицы для этих подгрупп признаков.

Таблица 3

Первая подгруппа признаков

№	Функция	Датчик	Интервал	Вес
1	$t_{0n} - t_{0л}$	РТМ	$> 0,53$	<b>3,4</b>
2	$t_{1n} - t_{1л}$	РТМ	(0,08; 0,45)	<b>6,1</b>
3	$t_{2n} - t_{2л}$	РТМ	$< 0,25$	<b>32,2</b>
4	$t_{3n} - t_{3л}$	РТМ	(0,25; 0,45)	<b>9,0</b>
5	$t_{4n} - t_{4л}$	РТМ	$< -0,11$	<b>10,4</b>
6	$t_{5n} - t_{5л}$	РТМ	$< -0,03$	<b>2,7</b>
7	$t_{6n} - t_{6л}$	РТМ	$< 0,08$	<b>16,1</b>
8	$t_{7n} - t_{7л}$	РТМ	$< 0,13$	<b>19,5</b>
9	$t_{8n} - t_{8л}$	РТМ	$< 0,08$	<b>15,3</b>
10	$t_{9n} - t_{9л}$	РТМ	$> 0,02$	<b>3,7</b>
11	$\frac{t_{0n} + t_{1n} + t_{2n}}{3} - \frac{t_{0л} + t_{1л} + t_{2л}}{3}$	РТМ	$< 0,21$	<b>14,6</b>
12	$\frac{t_{0n} + t_{2n} + t_{3n}}{3} - \frac{t_{0л} + t_{2л} + t_{3л}}{3}$	РТМ	$> 0,29$	<b>3,3</b>
13	$\frac{t_{0n} + t_{3n} + t_{4n}}{3} - \frac{t_{0л} + t_{3л} + t_{4л}}{3}$	РТМ	$> 0,34$	<b>4,2</b>
14	$\frac{t_{0n} + t_{4n} + t_{5n}}{3} - \frac{t_{0л} + t_{4л} + t_{5л}}{3}$	РТМ	(0; 0,31)	<b>11,4</b>
15	$\frac{t_{0n} + t_{5n} + t_{6n}}{3} - \frac{t_{0л} + t_{5л} + t_{6л}}{3}$	РТМ	$< 0,03$	<b>21,4</b>
16	$\frac{t_{0n} + t_{6n} + t_{7n}}{3} - \frac{t_{0л} + t_{6л} + t_{7л}}{3}$	РТМ	$< 0,09$	<b>32,5</b>
17	$\frac{t_{0n} + t_{7n} + t_{8n}}{3} - \frac{t_{0л} + t_{7л} + t_{8л}}{3}$	РТМ	$< 0,09$	<b>16,4</b>
18	$\frac{t_{0n} + t_{8n} + t_{1n}}{3} - \frac{t_{0л} + t_{8л} + t_{1л}}{3}$	РТМ	$> 0,33$	<b>26,2</b>
19	$t_{0n} - t_{0л}$	ИК	$> 0,01$	<b>13,0</b>

Таблица 4

## Вторая подгруппа признаков

№	Функция	Датчик	Интервал	Вес
20	$t_{7n} - t_{7л}$	ИК	(0,12; 0,63)	<b>5,1</b>
21	$t_{8n} - t_{8л}$	ИК	< 0,03	<b>4,7</b>
22	$t_{9n} - t_{9л}$	ИК	(0,17; 0,58)	<b>4,7</b>
23	$\frac{t_{0n} + t_{1n} + t_{2n}}{3} - \frac{t_{0л} + t_{1л} + t_{2л}}{3}$	ИК	> 0,41	<b>3,7</b>
24	$\frac{t_{0n} + t_{2n} + t_{3n}}{3} - \frac{t_{0л} + t_{2л} + t_{3л}}{3}$	ИК	< -0,01	<b>3,3</b>
25	$\frac{t_{0n} + t_{3n} + t_{4n}}{3} - \frac{t_{0л} + t_{3л} + t_{4л}}{3}$	ИК	(0,04; 0,48)	<b>5,1</b>
26	$\frac{t_{0n} + t_{4n} + t_{5n}}{3} - \frac{t_{0л} + t_{4л} + t_{5л}}{3}$	ИК	< 0,17	<b>4,8</b>
27	$\frac{t_{0n} + t_{5n} + t_{6n}}{3} - \frac{t_{0л} + t_{5л} + t_{6л}}{3}$	ИК	> 0,46	<b>3,5</b>
28	$\frac{t_{0n} + t_{6n} + t_{7n}}{3} - \frac{t_{0л} + t_{6л} + t_{7л}}{3}$	ИК	> 0,46	<b>8,2</b>
29	$\frac{t_{0n} + t_{7n} + t_{8n}}{3} - \frac{t_{0л} + t_{7л} + t_{8л}}{3}$	ИК	> 0,47	<b>9,3</b>
30	$\frac{t_{0n} + t_{8n} + t_{1n}}{3} - \frac{t_{0л} + t_{8л} + t_{1л}}{3}$	ИК	(0,09; 0,44)	<b>3,7</b>

Таблица 5

## Третья подгруппа признаков

№	Функция	Датчик	Интервал	Вес
31	$\frac{(t_{0,PTM} - t_{0,ИК})_n + (t_{6,PTM} - t_{6,ИК})_n + (t_{7,PTM} - t_{7,ИК})_n}{3} - \frac{(t_{0,PTM} - t_{0,ИК})_л + (t_{6,PTM} - t_{6,ИК})_л + (t_{7,PTM} - t_{7,ИК})_л}{3}$	–	> 0,28	<b>6,8</b>
32	$\frac{(t_{0,PTM} - t_{0,ИК})_n + (t_{7,PTM} - t_{7,ИК})_n + (t_{8,PTM} - t_{8,ИК})_n}{3} - \frac{(t_{0,PTM} - t_{0,ИК})_л + (t_{7,PTM} - t_{7,ИК})_л + (t_{8,PTM} - t_{8,ИК})_л}{3}$	–	< -0,01	<b>7,1</b>
33	$\frac{(t_{0,PTM} - t_{0,ИК})_n + (t_{8,PTM} - t_{8,ИК})_n + (t_{1,PTM} - t_{1,ИК})_н}{3} - \frac{(t_{0,PTM} - t_{0,ИК})_л + (t_{8,PTM} - t_{8,ИК})_л + (t_{1,PTM} - t_{1,ИК})_л}{3}$	–	< -0,01	<b>5,8</b>
34	$(\max_{i=1,8} t_i - \min_{i=1,8} t_i)_n - (\max_{i=1,8} t_i - \min_{i=1,8} t_i)_л$	РТМ	(0,08; 0,35)	<b>6,5</b>
35	$(\max_{i=1,8} t_i - \min_{i=1,8} t_i)_н - (\max_{i=1,8} t_i - \min_{i=1,8} t_i)_л$	ИК	< 0,11	<b>2,7</b>
36	$(\max_{i=1,8} t_i - \min_{i=1,8} t_i)_н - (\max_{i=1,8} t_i - \min_{i=1,8} t_i)_л$	разность РТМ и ИК	< 0,01	<b>2,7</b>
37	$\left( \sqrt{\sum_{i=1}^8 \frac{(t_i - t_{cp})^2}{8}} \right)_н - \left( \sqrt{\sum_{i=1}^8 \frac{(t_i - t_{cp})^2}{8}} \right)_л$	РТМ	> 0,11	<b>12,9</b>

Таблица 5

## Третья подгруппа признаков

№	Функция	Датчик	Интервал	Вес
38	$\left(\sqrt{\sum_{i=1}^8 \frac{(t_i - t_{cp})^2}{8}}\right)_n - \left(\sqrt{\sum_{i=1}^8 \frac{(t_i - t_{cp})^2}{8}}\right)_л$	ИК	(0,02; 0,18)	<b>2,9</b>
39	$\left(\sqrt{\sum_{i=1}^8 \frac{(t_i - t_{cp})^2}{8}}\right)_n - \left(\sqrt{\sum_{i=1}^8 \frac{(t_i - t_{cp})^2}{8}}\right)_л$	разность РТМ и ИК	> 0,19	<b>7,4</b>
40	$\sqrt{\sum_{i=0}^8 \frac{(\Delta t_i - t_{мжс})^2}{9}}$	РТМ	(0,12; 0,17)	<b>2,3</b>
41	$\sqrt{\sum_{i=0}^8 \frac{(\Delta t_i - t_{мжс})^2}{9}}$	ИК	> 0,43	<b>19,1</b>
42	$\sqrt{\sum_{i=0}^8 \frac{(\Delta t_i - t_{мжс})^2}{9}}$	разность РТМ и ИК	(0,25; 0,44)	<b>31,2</b>

Таблица 6

## Матрица парных коэффициентов корреляции

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	1,0	0,2	0,4	0,3	0,3	0,4	0,4	0,3	0,3	0,0	<b>0,8</b>	<b>0,7</b>	<b>0,8</b>						
2	0,2	1,0	0,3	0,2	0,2	0,1	0,3	0,3	0,4	0,1	<b>0,7</b>	0,3	0,3	0,2	0,3	0,4	0,4	<b>0,7</b>	0,2
3	0,4	0,3	1,0	0,4	0,1	0,3	0,3	0,2	0,2	-0,2	<b>0,7</b>	<b>0,8</b>	0,4	0,4	0,5	0,4	0,4	0,4	0,4
4	0,3	0,2	0,4	1,0	0,5	0,5	0,3	0,3	0,2	-0,2	0,4	<b>0,7</b>	<b>0,7</b>	0,5	0,5	0,4	0,4	0,3	0,3
5	0,3	0,2	0,1	0,5	1,0	0,5	0,0	0,1	0,2	-0,1	0,3	0,4	<b>0,8</b>	<b>0,7</b>	0,4	0,2	0,3	0,3	0,3
6	0,4	0,1	0,3	0,5	0,4	1,0	0,4	0,1	0,2	-0,1	0,3	0,5	0,6	<b>0,8</b>	<b>0,7</b>	0,4	0,3	0,3	0,3
7	0,4	0,3	0,3	0,3	0,0	0,3	1,0	0,4	0,2	-0,1	0,4	0,5	0,3	0,4	<b>0,7</b>	<b>0,8</b>	0,4	0,4	0,4
8	0,3	0,3	0,2	0,3	0,1	0,1	0,4	1,0	0,4	-0,1	0,4	0,4	0,4	0,3	0,4	<b>0,7</b>	<b>0,7</b>	0,5	0,3
9	0,3	0,4	0,2	0,2	0,2	0,2	0,2	0,4	1,0	0,0	0,4	0,3	0,3	0,3	0,3	0,4	<b>0,7</b>	<b>0,8</b>	0,2
10	0,0	0,1	-0,2	-0,2	-0,1	-0,1	-0,1	-0,1	0,0	1,0	0,0	-0,1	-0,1	0,0	-0,1	0,0	0,0	0,1	0,0
11	<b>0,8</b>	<b>0,7</b>	<b>0,7</b>	0,4	0,3	0,3	0,4	0,4	0,4	0,0	1,0	<b>0,9</b>	<b>0,7</b>	<b>0,7</b>	<b>0,7</b>	<b>0,7</b>	<b>0,7</b>	<b>0,9</b>	0,6
12	<b>0,8</b>	0,3	<b>0,8</b>	<b>0,7</b>	0,4	0,5	0,5	0,4	0,3	-0,1	<b>0,9</b>	1,0	<b>0,9</b>	<b>0,8</b>	<b>0,8</b>	<b>0,8</b>	<b>0,7</b>	<b>0,7</b>	<b>0,7</b>
13	<b>0,8</b>	0,3	0,4	<b>0,7</b>	<b>0,8</b>	0,6	0,3	0,4	0,3	-0,1	<b>0,7</b>	<b>0,9</b>	1,0	<b>0,9</b>	<b>0,8</b>	<b>0,7</b>	<b>0,7</b>	0,6	<b>0,7</b>
14	<b>0,8</b>	0,2	0,4	0,5	<b>0,7</b>	<b>0,8</b>	0,4	0,3	0,3	0,0	<b>0,7</b>	<b>0,8</b>	<b>0,9</b>	1,0	<b>0,9</b>	<b>0,7</b>	0,6	0,6	0,6
15	<b>0,8</b>	0,3	0,5	0,5	0,4	<b>0,7</b>	<b>0,7</b>	0,4	0,3	-0,1	<b>0,7</b>	<b>0,8</b>	<b>0,8</b>	<b>0,9</b>	1,0	<b>0,9</b>	<b>0,7</b>	0,6	<b>0,7</b>
16	<b>0,8</b>	0,4	0,4	0,4	0,2	0,4	<b>0,8</b>	<b>0,7</b>	0,4	0,0	<b>0,7</b>	<b>0,8</b>	<b>0,7</b>	<b>0,7</b>	<b>0,9</b>	1,0	<b>0,8</b>	<b>0,7</b>	<b>0,7</b>
17	<b>0,8</b>	0,4	0,4	0,4	0,3	0,3	0,4	<b>0,7</b>	<b>0,7</b>	0,0	<b>0,7</b>	<b>0,7</b>	<b>0,7</b>	0,6	<b>0,7</b>	<b>0,8</b>	1,0	<b>0,9</b>	0,6
18	<b>0,7</b>	<b>0,7</b>	0,4	0,3	0,3	0,3	0,4	0,5	<b>0,8</b>	0,1	<b>0,9</b>	<b>0,7</b>	0,6	0,6	0,6	<b>0,7</b>	<b>0,9</b>	1,0	0,6
19	<b>0,8</b>	0,2	0,4	0,3	0,3	0,3	0,4	0,3	0,2	0,0	0,6	<b>0,7</b>	<b>0,7</b>	0,6	<b>0,7</b>	<b>0,7</b>	0,6	0,6	1,0

Таблица 7

Матрица парных коэффициентов корреляции

	20	21	22	23	24	25	26	27	28	29	30
20	1,0	0,3	0,0	0,4	0,3	0,3	0,2	0,3	0,7	0,7	0,5
21	0,3	1,0	0,1	0,2	0,1	0,2	0,1	0,2	0,3	0,7	0,7
22	0,0	0,1	1,0	0,0	-0,1	0,0	0,0	0,2	0,1	0,0	0,1
23	0,4	0,2	0,0	1,0	0,7	0,5	0,4	0,4	0,6	0,5	0,7
24	0,3	0,1	-0,1	0,7	1,0	0,8	0,5	0,5	0,5	0,4	0,4
25	0,3	0,2	0,0	0,5	0,8	1,0	0,8	0,7	0,6	0,5	0,5
26	0,2	0,1	0,0	0,4	0,5	0,8	1,0	0,8	0,6	0,5	0,4
27	0,3	0,2	0,2	0,4	0,5	0,7	0,8	1,0	0,8	0,6	0,5
28	0,7	0,3	0,1	0,6	0,5	0,6	0,6	0,8	1,0	0,8	0,7
29	0,7	0,7	0,0	0,5	0,4	0,5	0,5	0,6	0,8	1,0	0,9
30	0,5	0,7	0,1	0,7	0,4	0,5	0,4	0,5	0,7	0,9	1,0

Таблица 8

Матрица парных коэффициентов корреляции

	31	32	33	34	35	36	37	38	39	40	41	42
31	1,0	0,7	0,4	0,1	0,3	0,3	0,1	0,3	0,3	-0,1	-0,2	-0,1
32	0,7	1,0	0,7	0,1	0,2	0,2	0,1	0,3	0,3	-0,1	0,0	0,0
33	0,4	0,7	1,0	0,0	0,1	-0,1	0,0	0,1	0,0	-0,1	-0,1	-0,1
34	0,1	0,1	0,0	1,0	0,6	0,3	0,9	0,6	0,3	0,0	0,3	0,2
35	0,3	0,2	0,1	0,6	1,0	0,8	0,5	0,9	0,8	0,1	0,3	0,3
36	0,3	0,2	-0,1	0,3	0,8	1,0	0,3	0,8	1,0	0,0	0,3	0,3
37	0,1	0,1	0,0	0,9	0,5	0,3	1,0	0,6	0,3	0,0	0,2	0,1
38	0,3	0,3	0,1	0,6	0,9	0,8	0,6	1,0	0,8	0,1	0,3	0,3
39	0,3	0,3	0,0	0,3	0,8	1,0	0,3	0,8	1,0	0,0	0,3	0,3
40	-0,1	-0,1	-0,1	0,0	0,1	0,0	0,0	0,1	0,0	1,0	0,4	0,0
41	-0,2	0,0	-0,1	0,3	0,3	0,3	0,2	0,3	0,3	0,4	1,0	0,7
42	-0,1	0,0	-0,1	0,2	0,3	0,3	0,1	0,3	0,3	0,0	0,7	1,0

Кроме того, в первой группе были выделены признаки, которые слабо коррелируют с другими признаками в данной группе, – это признаки под номерами 10, 22 и 40 (см. табл. 6, 7, 8).

Аналогичная картина наблюдалась и для других групп диагностических признаков.

Ранее для оценки диагностической способности каждого признака была применена величина, которую традиционно относят к информативностям, а именно статистическая информативность:

$$I_l^q = -\ln \left( \frac{C_M^m C_K^k}{C_{M+K}^{m+k}} \right), \quad (3)$$

где  $M$  – количество молочных желез класса «Здоровые»,  $m$  – количество молочных желез класса «Здоровые», у которых выполнен признак;  $K$  – количество молочных желез класса «Больные»,  $k$  – количество молочных желез класса «Больные», у которых выполнен признак.

Таким образом, в каждой выделенной подгруппе признаков (см. табл. 3, 4, 5) целесообразно оптимизировать число диагностических признаков по следующему правилу: из каждой пары признаков с коэффициентом  $\rho_g \geq 0,7$  оставляем признак с большей информативностью; если же для пары признаков  $\rho_g < 0,7$ , то оставляем оба признака. Начинать проверку надо с признаков с наименьшей информативностью.

Ниже представлены таблицы отобранных высокоинформативных признаков из первой группы зависимых признаков (табл. 9, 10, 11).

Таблица 9

**Признаки первой подгруппы**

№	Функция	Датчик	Интервал	Вес
3	$t_{2n} - t_{2,l}$	РТМ	< 0,25	<b>32,2</b>
4	$t_{3n} - t_{3,l}$	РТМ	(0,25; 0,45)	<b>9,0</b>
10	$t_{9n} - t_{9,l}$	РТМ	> 0,02	<b>3,7</b>
14	$\frac{t_{0n} + t_{4n} + t_{5n}}{3} - \frac{t_{0,l} + t_{4,l} + t_{5,l}}{3}$	РТМ	(0; 0,31)	<b>11,4</b>
16	$\frac{t_{0n} + t_{6n} + t_{7n}}{3} - \frac{t_{0,l} + t_{6,l} + t_{7,l}}{3}$	РТМ	< 0,09	<b>32,5</b>
18	$\frac{t_{0n} + t_{8n} + t_{1n}}{3} - \frac{t_{0,l} + t_{8,l} + t_{1,l}}{3}$	РТМ	> 0,33	<b>26,2</b>

Таблица 10

**Признаки второй подгруппы**

№	Функция	Датчик	Интервал	Вес
22	$t_{9n} - t_{9,l}$	ИК	(0,17; 0,58)	<b>4,7</b>
23	$\frac{t_{0n} + t_{1n} + t_{2n}}{3} - \frac{t_{0,l} + t_{1,l} + t_{2,l}}{3}$	ИК	> 0,41	<b>3,7</b>
25	$\frac{t_{0n} + t_{3n} + t_{4n}}{3} - \frac{t_{0,l} + t_{3,l} + t_{4,l}}{3}$	ИК	(0,04; 0,48)	<b>5,1</b>
29	$\frac{t_{0n} + t_{7n} + t_{8n}}{3} - \frac{t_{0,l} + t_{7,l} + t_{8,l}}{3}$	ИК	> 0,47	<b>9,3</b>
30	$\frac{t_{0n} + t_{8n} + t_{1n}}{3} - \frac{t_{0,l} + t_{8,l} + t_{1,l}}{3}$	ИК	(0,09; 0,44)	<b>3,7</b>

Таблица 11

**Признаки третьей подгруппы**

№	Функция	Датчик	Интервал	Вес
32	$\frac{(t_{0,РТМ} - t_{0,ИК})_n + (t_{7,РТМ} - t_{7,ИК})_n + (t_{8,РТМ} - t_{8,ИК})_n}{3} - \frac{(t_{0,РТМ} - t_{0,ИК})_l + (t_{7,РТМ} - t_{7,ИК})_l + (t_{8,РТМ} - t_{8,ИК})_l}{3}$	-	< -0,01	<b>7,1</b>
37	$\left( \sqrt{\sum_{i=1}^8 \frac{(t_i - t_{cp})^2}{8}} \right)_n - \left( \sqrt{\sum_{i=1}^8 \frac{(t_i - t_{cp})^2}{8}} \right)_l$	РТМ	> 0,11	<b>12,9</b>
39	$\left( \sqrt{\sum_{i=1}^8 \frac{(t_i - t_{cp})^2}{8}} \right)_n - \left( \sqrt{\sum_{i=1}^8 \frac{(t_i - t_{cp})^2}{8}} \right)_l$	разность РТМ и ИК	> 0,19	<b>7,4</b>
40	$\sqrt{\sum_{i=0}^8 \frac{(\Delta t_i - \overline{t_{мж}})^2}{9}}$	РТМ	(0,12; 0,17)	<b>2,3</b>
42	$\sqrt{\sum_{i=0}^8 \frac{(\Delta t_i - \overline{t_{мж}})^2}{9}}$	разность РТМ и ИК	(0,25; 0,44)	<b>31,2</b>

Пользуясь случаем, авторы выражают благодарность участнику семинара «Математическое моделирование в медицине» Владиславу Викторовичу Левшинскому.

### ПРИМЕЧАНИЕ

<sup>1</sup> Работа выполнена при финансовой поддержке РФФИ и АВО (проект № 15-47-02475-р\_поволжье\_a).

### СПИСОК ЛИТЕРАТУРЫ

1. К вопросу об оптимизации скринингового обследования молочных желез методом микроволновой радиотермометрии / Т. В. Замечник, Е. А. Мазепа, С. И. Черкесова, Г. В. Панкова // Вестник новых медицинских технологий. – 2014. – Т. 21, № 4. – С. 34–38.
2. Лосев, А. Г. О взаимосвязи некоторых признаков РТМ-диагностики заболеваний молочных желез / А. Г. Лосев, Е. А. Мазепа, Х. М. Сулейманова // Вестник Волгоградского государственного университета. Серия 1, Математика. Физика. – 2015. – № 4 (29). – С. 35–44.
3. Лосев, А. Г. О некоторых характерных признаках в диагностике патологии молочных желез по данным микроволновой радиотермометрии / А. Г. Лосев, Е. А. Мазепа, Т. В. Замечник // Современные проблемы науки и образования. – 2014. – № 6. – С. 254.
4. Медик, В. А. Статистика в медицине и биологии. Руководство. В 2 т. Т. 1. Теоретическая статистика / В. А. Медик, М. С. Токмачев, Б. Б. Фишман. – М. : Медицина, 2000. – 412 с.
5. Моисеенко, В. М. Кинетические особенности роста рака молочной железы и их значение для раннего выявления опухоли / В. М. Моисеенко, В. Ф. Семиглазов // Маммология. – 1997. – № 3. – С. 311.
6. Применение радиотермометрии для диагностики рака молочной железы / Л. М. Бурдина [и др.] // Маммология. – 1998. – № 2. – С. 312.
7. Тихомирова, Н. Н. Техника проведения РТМ-обследования молочных желез / Н. Н. Тихомирова. – 2008. – 64 с.

### REFERENCES

1. Zamechnik T.V., Mazepa E.A., Cherkesova S.I., Pankova G.V. K voprosu ob optimizatsii skrininogovogo obsledovaniya molochnykh zhelez metodom mikrovolnovoy radiotermometrii [The Question about Optimization of Screening of the Mammary Glands by the Method of Microwave Radiometry]. *Vestnik novykh meditsinskikh tekhnologiy* [Science Journal of New Medical Technology], 2014. vol. 21, iss. 4, pp. 34-38.
2. Losev A.G., Mazepa E.A., Suleymanova Kh.M. O vzaimosvyazi nekotorykh priznakov RTM-diagnostiki zabolevaniy molochnykh zhelez [On the Relationship of Certain Characteristics of RTM-Diagnosis of Breast Diseases]. *Vestnik Volgogradskogo gosudarstvennogo universiteta. Seriya 1, Matematika. Fizika* [Science Journal of Volgograd State University. Mathematics. Physics], 2015, no. 4 (29), pp. 35-44.
3. Losev A.G., Mazepa E.A., Zamechnik T.V. O nekotorykh kharakternykh priznakakh v diagnostike patologii molochnykh zhelez po dannym mikrovolnovoy radiotermometrii [About Some Characteristic Symptoms in the Diagnosis of Breast Pathology on the Data of Microwave Radiometry]. *Sovremennye problemy nauki i obrazovaniya*, 2014, no. 6, p. 254.
4. Medik V.A., Tokmachev M.S., Fishman B.B. *Statistika v meditsine i biologii. Rukovodstvo. V 2 t. T. 1. Teoreticheskaya statistika* [Statistics in Medicine and Biology. In 2 vols. Vol. 1. Theoretical Statistics]. Moscow, Meditsina Publ., 2000. 412 p.
5. Moiseenko V.M., Semiglazov V.F. Kineticheskie osobennosti rosta raka molochnoy zhelezy i ikh znachenie dlya rannego vyyavleniya opukholi [Kinetic Features of Breast Cancer Growth and Their Importance for Early Cancer Detection]. *Mammologiya*, 1997, no. 3, pp. 3-11.
6. Burdina L.M., et al. Primenenie radiotermometrii dlya diagnostiki raka molochnoy zhelezy [Application of Microwave Radiometry for Detection of Breast Cancer]. *Mammologiya*, 1998, no. 2, pp. 3-12.
7. Tikhomirova N.N. *Tekhnika provedeniya RTM-obsledovaniya molochnykh zhelez* [The Technique of the RTM-Breast Examination]. 2008. 64 p.

## ON OPTIMIZATION OF THE NUMBER OF DIAGNOSTIC SIGNS FOR BREAST DISEASES THROUGH THERMOMETRIC DATA

**Elena Alekseevna Mazepa**

Candidate of Physical and Mathematical Sciences, Associate Professor,  
Department of Mathematical Analysis and Function Theory,  
Volgograd State University  
lmazepa@rambler.ru, matf@volsu.ru  
Prosp. Universitetsky, 100, 400062 Volgograd, Russian Federation

**Khedi Movladovna Suleymanova**

Postgraduate Student, Department of Mathematical Analysis and Function Theory,  
Volgograd State University  
hedi.medi@mail.ru, matf@volsu.ru  
Prosp. Universitetsky, 100, 400062 Volgograd, Russian Federation

**Abstract.** The work is devoted, on the one hand, to the study of the correlation interconnection of signs, derived from medical thermometer data intended for the diagnosis of breast diseases. The authors also highlighted the most important signs of a disease for the express-diagnostics of cancer breast tumors.

One of the directions of development of artificial intelligence systems is the development of expert systems for medical diagnostics. Their use helps the doctor to improve the quality of their work. The objective of such systems is not only the definition but also consulting assistance in identifying diseases (one or more) of patients, based on his observations. When creating intelligent advisory systems, developers rely on a number of high-quality diagnostic features. For their mathematical interpretation, a variety of functional relationships between the initial thermometric data can be used. In this regard, the number of possible diagnostic quantitative traits that may be employed in diagnostic systems express increases to several hundred or even thousands. Each of the resulting symptoms was assessed in terms of its information content, which reflects the degree of diagnostic ability of this feature. However, many of the quantified traits are interdependent and interchangeable. Reducing the number of diagnostic features with regard to their independence and maximum information content is one of the ways of increasing the quality of advisory intelligent systems.

**Key words:** microwave radiometry, intellectual consulting systems, express-diagnostics, breast tumors, correlation analysis.