Хирургия. Журнал имени Н.И. Пирогова
2024, №9, с. 76–85
https://doi.org/10.17116/hirurgia202409176

Pirogov Russian Journal of Surgery
2024, No. 9, pp. 76–85
https://doi.org/10.17116/hirurgia202409176

Прогноз рецидива при папиллярном раке щитовидной железы по дооперационным данным

© Н.С. КУЗНЕЦОВ[1], М.В. СКИБИЦКАЯ[1], А.П. ВАЙНШТОК[2], Е.А. ВАЩЕНКО[2]

[1]ГНЦ РФ ФГБУ «Национальный медицинский исследовательский центр эндокринологии» Минздрава России, Москва, Россия;
[2]ФГБУН «Институт проблем передачи информации им. А.А. Харкевича» Российской академии наук, Москва, Россия

Резюме

Цель исследования. Создание формализованного метода прогноза рецидива папиллярного рака щитовидной железы (ПРЩЖ) после гемитиреоидэктомии по данным дооперационного обследования.

Материал и методы. На данном этапе исследования из БД Эндокринологического научного центра выбран 101 пациент с диагнозом ПРЩЖ, проходивших оперативное лечение в хирургическом отделении ГНЦ РФ ФГБУ «Национальный медицинский исследовательский центр эндокринологии» Минздрава России в 2017—2023 гг. У 47 пациентов выявлен рецидив, у 54 больных рецидива не было в течение 5 лет после первичного оперативного лечения, т. е. в эту группу отобраны пациенты, у которых операция была в 2017—2018 гг.

Для нахождения правил прогноза применен оригинальный метод классификации, разработанный в Институте проблем передачи информации им. А.А. Харкевича РАН, основанный на поиске подмножеств показателей (признаков) и построении кусочно-линейных правил, попарно разделяющих классы, и последующем голосовании таких правил для принятия решения.

Результаты. Выполнен экзамен на учебной выборке (101 прецедент) и методом скользящего контроля (10 испытаний на 10 случайно отобранных прецедентах). На учебной выборке чувствительность прогностического алгоритма достигает 91% при специфичности 78% и вероятности ошибки 13%. Агрегированный результат 10 испытаний методом скользящего контроля показал чувствительность прогностического алгоритма 86% при специфичности 75% и вероятности ошибки 15%. Такой результат близок к результатам на полной выборке, что подтверждает эффективность используемого метода для прогноза рецидива.

Заключение. Выполненные пилотные эксперименты показали наличие закономерностей в данных, позволяющих решать задачу прогноза рецидива по показателям дооперационного обследования, и перспективность применяемого метода. Дальнейшее изучение проблемы может помочь врачам в принятии решений, а также внести значительные коррективы в тактику ведения пациентов с ПРЩЖ.

Ключевые слова: *щитовидная железа, папиллярный рак, дооперационное обследование хирургическое лечение, рецидив, прогноз.*

Информация об авторах:
Кузнецов Н.С. — https://orcid.org/0000-0002-9419-7013
Скибицкая М.В. —e-mail: Skibitskaya_93@mail.ru; https://orcid.org/0000-0003-1908-1720
Вайншток А.П. — e-mail: wein@iitp.ru; https://orcid.org/0000-0001-5032-1038
Ващенко Е.А. — e-mail: Vea@iitp.ru
Автор, ответственный за переписку: Скибицкая М.В. — e-mail: skibitskaya_93@mail.ru

Как цитировать:
Кузнецов Н.С., Скибицкая М.В., Вайншток А.П., Ващенко Е.А. Прогноз рецидива при папиллярном раке щитовидной железы по дооперационным данным. *Хирургия. Журнал им. Н.И. Пирогова*. 2024;9:76–85. https://doi.org/10.17116/hirurgia202409176

Prediction of papillary thyroid cancer recurrence according to preoperative data

© N.S. KUZNETSOV[1], M.V. SKIBITSKAYA[1], A.P. VAYNSHTOK[2], E.A. VASHCHENKO[2]

[1]Endocrinology Research Center, Moscow, Russia;
[2] Kharkevich Institute for Information Transmission Problems, Moscow, Russia

Abstract

Objective. To create a formalized method for predicting papillary thyroid cancer recurrence after hemithyroidectomy based on preoperative data.

Material and methods. At this stage of the study, we selected 101 patients with papillary thyroid cancer who underwent surgical treatment in 2017—2023. Recurrence was observed in in 47 patients. Fifty-four patients had no recurrence within 5 years after surgical treatment, i.e. these patients underwent surgery in 2017—2018. To find prediction rules, we used original classification method based on searching for subsets of variables and piecewise linear rules separating classes in pairs with subsequent voting of such rules to make a decision.

Results. The exam was carried out using a training sample (101 cases) and sliding control method (10 tests on 10 random cases). On the training sample, sensitivity of predictive algorithm was 91%, specificity 78% and error rate 13%. The aggregated result of 10 trials using sliding control method revealed sensitivity of predictive algorithm 86%, specificity 75% and error rate 15%. This result is close to overall sample and confirms the effectiveness of this method for predicting recurrence.

Conclusion. The pilot experiments revealed the patterns in data for potential prediction of recurrence based on preoperative indicators. Further study of this problem may be valuable for decision-making and adjustments in the management of patients with papillary thyroid cancer.

Keywords: *thyroid gland, papillary cancer, preoperative examination, surgical treatment, recurrence, prognosis.*

Information about the authors:
Kuznetsov N.S. — https://orcid.org/0000-0002-9419-7013
Skibitskaya M.V. —e-Email: Skibitskaya_93@mail.ru; https://orcid.org/0000-0003-1908-1720
Vaynshtok A.P. — e-mail: wein@iitp.ru; https://orcid.org/0000-0001-5032-1038
Vashchenko E.A. — e-mail: Vea@iitp.ru
Corresponding author: Skibitskaya M.V. — e-mail: skibitskaya_93@mail.ru

To cite this article:
Kuznetsov NS, Skibitskaya MV, Vaynshtok AP, Vashchenko EA. Prediction of papillary thyroid cancer recurrence according to preoperative data. *Pirogov Russian Journal of Surgery*. 2024;9:76–85. (In Russ.). https://doi.org/10.17116/hirurgia202409176

Введение

Современной задачей и тенденцией в медицине является персонализация лечебного процесса. Это требует анализа индивидуальных факторов больного, своевременной диагностики и выбора оптимальной тактики лечения [1]. Папиллярный рак щитовидной железы (ПРЩЖ) является наиболее распространенным эндокринным злокачественным заболеванием [2, 3]. Он составляет примерно 85% [4] среди дифференцированного рака щитовидной железы. Несмотря на то что ПРЩЖ хорошо дифференцирован и имеет низкую частоту местной инвазии, метастазы опухоли в лимфатические узлы (ЛУ) встречаются довольно часто [5]. Примерно у половины пациентов, у которых клинически диагностирован папиллярный рак, выявляется рецидив в течение 5 лет после первичного хирургического лечения [6]. Кроме того, выделяют гистологические варианты более агрессивного течения с отчетливыми клиническими, патологическими и молекулярными особенностями. К наиболее агрессивным вариантам относятся ПРЩЖ из высоких клеток, ПРЩЖ из столбчатых клеток, ПРЩЖ из клеток типа «сапожных гвоздей» (hobnail variant), а также диффузно-склерозирующий вариант ПРЩЖ. На его долю приходится около 5% [7, 8] всех случаев РЩЖ. Эти случаи имеют более высокую частоту рецидива и метастазирования, а также в некоторых наблюдениях отсутствует чувствительность к терапии радиоактивным йодом (RAI) [9].

Материал и методы

Эффективная тактика лечения ПРЩЖ учитывает стадию заболевания, которая устанавливается в процессе обследования. Если осуществляется хирургическое лечение, то стадия заболевания уточняется после операции (морфологического исследования). Предоперационное суждение (дооперационное стадирование, врачебный прогноз) о фактической стадии заболевания является приближенной оценкой послеоперационного стадирования и основывается на данных дооперационной диагностики (ультразвуковое исследование (УЗИ) щитовидной железы и региональных ЛУ, тонкоигольной аспирационной биопсии щитовидной железы — ТАБ). Различия в оценках до- и послеоперационного стадирования рассматриваются как диагностические ошибки. Эти ошибки, по данным работ [10, 11], встречаются в 25—40% случаев: гипердиагностика выявляется в 52—75% случаев, а гиподиагностика — в 20—25% случаев [11]. Наличие ошибок ведет либо к ухудшению результатов лечения больных (в случае недооценки стадии опухолевого процесса), либо к ухудшению качества жизни пациентов (в случае неадекватного объема оперативного вмешательства — 38,7%) [10].

Применяются две тактики хирургического лечения. Стандартной операцией при ПРЩЖ является экстрафасциальная тиреоидэктомия. Данный тип операции снижает риск летальности и прогрессирования заболевания, а также сопряжен с минимальным риском осложнений [12]. Гемитиреоидэктомию обычно выполняют у пациентов с папиллярной карциномой щитовидной железы низкого и среднего риска. Кроме того, по данным некоторых авторов [13], профилактическая центральная лимфодиссекция шеи значительно снижает риск рецидива в центральном отделе у пациентов, перенесших гемитиреоидэктомию.

Для выбора адекватной тактики лечения предпочтительно установить истинную стадию заболевания еще до операции. При выставлении стадии при

ПРЩЖ опираются на результаты клинического, инструментального, гистологического и лабораторного обследования больного, а также на наличие или отсутствие отдаленных метастазов в региональные ЛУ [12].

Точность УЗ-диагностики в выявлении ЛУ различна, она во многом зависит от локализации ЛУ, качества аппаратуры, а также опыта исследователя [14, 15]. УЗ-характеристики ЛУ, включая увеличение объема, изменения формы, нечеткие границы, внутреннюю кальцификацию и некроз, имеют диагностическую ценность при дифференциации доброкачественных и злокачественных узлов [16, 17].

Таким образом, разработка методов прогнозирования фактической стадии ПРЩЖ и возможности рецидива на основе комплекса дооперационных показателей является актуальной проблемой при выборе тактики лечения заболевания в дооперационный период. Частным случаем этой проблемы является задача прогнозирования агрессивного статуса опухолевого процесса у больных с ПРЩЖ.

Материал исследования (выборка)

В работе использованы данные пациентов, представленные в историях болезни из базы данных ГНЦ РФ ФГБУ «Национальный медицинский исследовательский центр эндокринологии» Минздрава России. Материал содержит сведения о 1002 больных за период с 2017 по 2023 г. с диагнозом злокачественного новообразования щитовидной железы (ПРЩЖ) и проходивших оперативное лечение в хирургическом отделении НМИЦ эндокринологии. Из этих пациентов на данном этапе исследования выбран 101 пациент, из которых у 47 выявлен рецидив, и 54 пациента без рецидива в течение 5 лет после первичного оперативного лечения. На **рис. 1** представлены гистограммы распределения пациентов по полу и возрасту. Средний возраст пациентов составил 45,7 лет.

Данные о пациентах включали: пол, возраст, дата первичной диагностики, дата первичной операции, УЗ-характеристики первичной опухоли (размер, контур, эхогенность, кальцинаты), количество метастазов; УЗ-характеристики ЛУ (количество пораженных ЛУ, контур, эхоструктура); уровень кальцитонина в крови, данные в соответствии с международной классификацией BETESDA (стандартизация цитологических критериев результатов тонкоигольной аспирационной биопсии узлов щитовидной железы); данные системы TIRADS для унифицированной оценки узловых образований щитовидной железы [18]; наличие микрокальцинатов; преобладание толщины узла над шириной (в УЗ-диагностике это соответствует термину «вертикальная пространственная ориентация узла» [19]); неровные контуры и выраженная гипоэхогенность [18, 19]; данные международной системы классификации TNM, описывающей распространение опухолевого процесса в щитовидной железе (T1—T4), вовлечение в процесс регионарных ЛУ (N0/N+) и отсутствие/наличие отдаленных метастазов (M0/M1); данные о послеоперационных осложнениях. Данные о наличии или отсутствии рецидива являлись целевым фактором.

Для ввода показателей пациентов в ЭВМ была применена специальная интерактивная программа, которая позволяет вводить числовые, текстовые и отдельные значения списочных данных. На выходе программы генерируется таблица в формате Excel, строками которой являются пациенты (объекты), а столбцами значения закодированных показателей и номер класса, к которому относится пациент (нет рецидива — 1, есть рецидив — 2). Эта таблица представляет конечный перечень объектов, для которых известно, к какому из классов они принадлежат и исходную информацию о них для решения задачи обучения классификации.

Для управления вводом данных применятся специальный язык разметки текста, который позволяет различать типы данных и осуществлять логические переходы, в зависимости от значения заглавного атрибута группы показателей, например, показатель «УЗИ ЛУ — патология» может принимать 2 значения: «патологии нет/патология есть». В случае если «па-

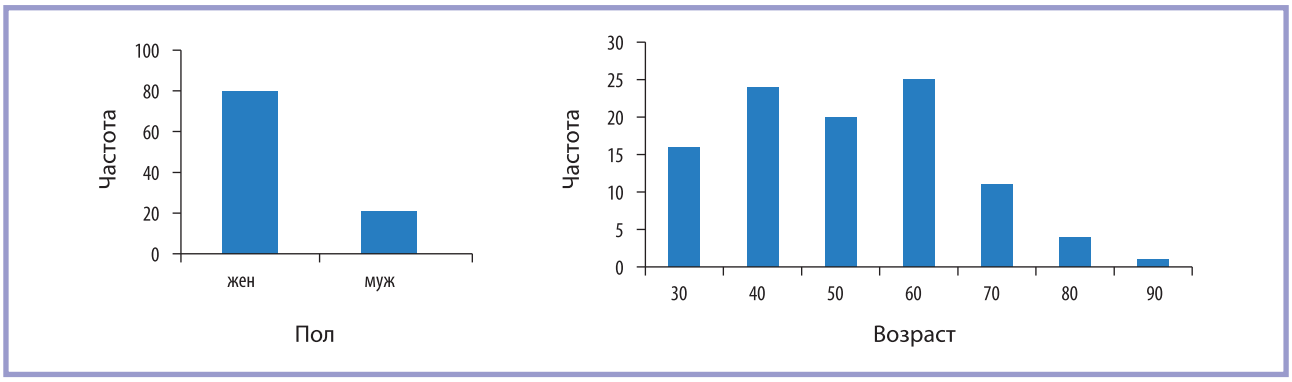

Рис. 1. **Гистограммы распределения пациентов по полу и возрасту.**
Fig. 1. **Gender- and age-adjusted distribution of patients.**

тологии нет» при вводе автоматически пропускается запрос всех показателей, наблюдаемых при наличии патологии.

На **рис. 2** представлена программа ввода: в левом окне осуществляется ввод значений показателей, а в правом отображаются введенные показатели, что позволяет проверить и исправить ошибки ввода непосредственно перед записью данных пациента в таблицу.

Полный список введенных в ЭВМ показателей, их идентификаторы и кодировка приведены в **приложении 1**.

В **табл. 1** представлен фрагмент Excel-таблицы закодированных показателей (из-за размера таблицы она разделена в тексте на 2 части). Показатели дооперационного обследования pol — triads и reced используются для нахождения правил прогноза рецидива. Показатели post_oper — ipsi представляют

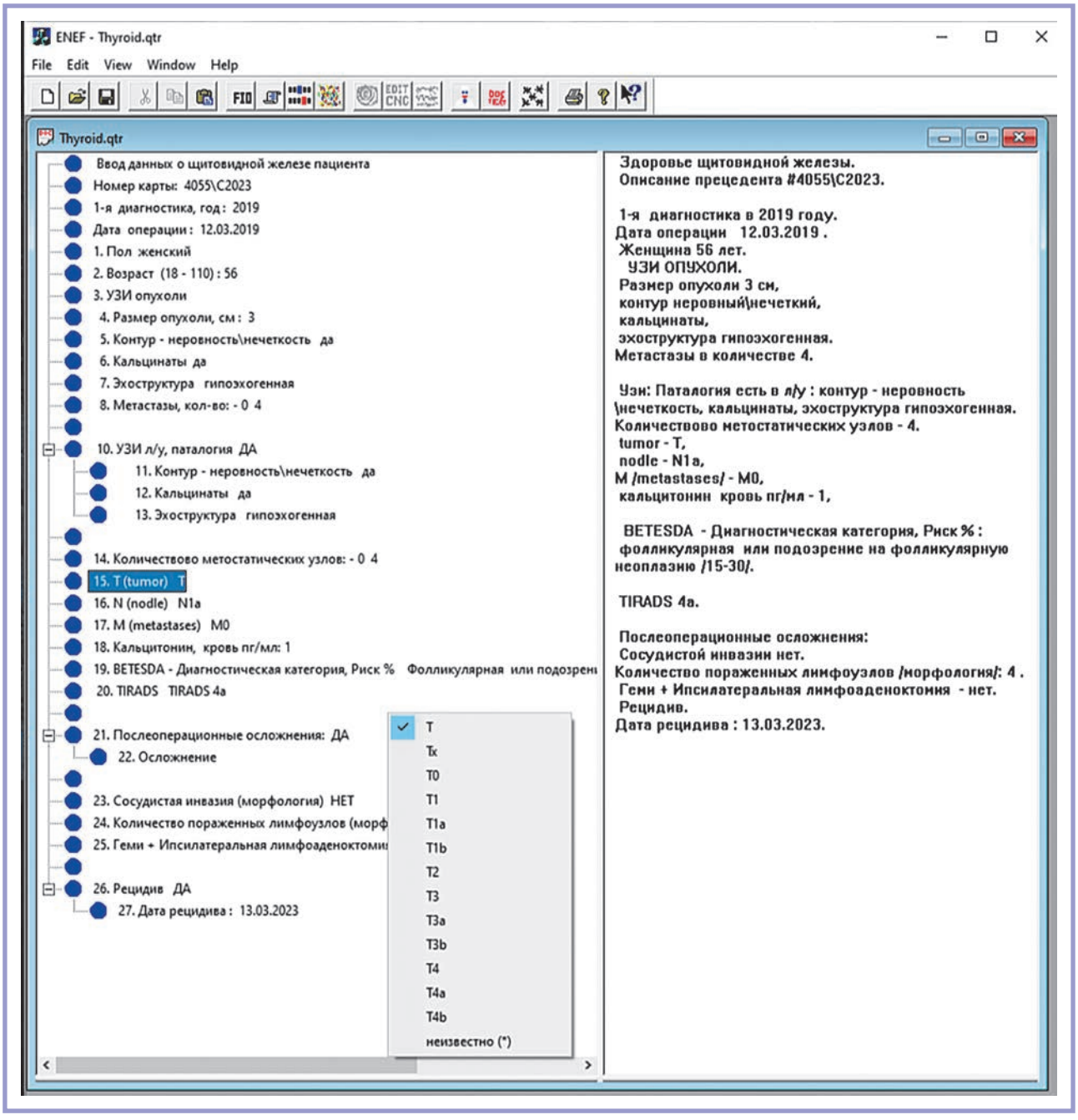

Рис. 2. **Программа ввода показателей.**
В левом окне осуществляется ввод значений показателей, а в правом — отображаются введенные показатели.

Fig. 2. **Program for entering indicators.**
In the left window, you enter the values of indicators, and in the right window the entered indicators are displayed.

Таблица 1. **Фрагмент таблицы закодированных показателей (результат программы ввода)**
Table 1. **Fragment of table with coded indicators (input program result)**

id	nun	yearl	oper_dat	Pol	age	size	cant	Ca	eho	mets	uzi_ nodes	Cont_ nodes	Ca_ nodes	eho_ nodes
1	65780\A21	2021	2022	2	54	Z.D	1	2	3	10	2	2	2	4
2	1040\C2021	2020	03.02.2021	1	27	3,0	2	1	3	5	2	1	1	4
3	12116\c2021	2021	01.09.2021	1	43	3,0	2	2	1	5	2	2	1	4
4	7985\c2022	2021	11.02.2022	1	57	5,0	1	1	3	2	2	2	2	4
5	8591\C2022	2021	anp.21	1	27	3,0	2	1	3	1	2	2	1	4

id	nun	tumor	node	n_mtst	ca_bl	betesda	tirads	post_oper	postoper	sosud_omg	qtymorf	ipsi	raced	data_rec
1	657S0\A21	7	7	1	6	3	7	1	1	1	9	1	2	2023
2	1ОМДС2021	7	1	1	1	5	4	1	1	1	5	1	1	1
3		7	7	1	5	3	6	1	1	1	5	1	2	фев. 22
4	79S5\c2O22	ID	1	3	1	5	6	2	3	2	3	1	2	2022
5	S591\C2O22	5	7	1	2	5	5	1	1	1	1	1	2	июн. 21

послеоперационное обследование и будут применяться в последующих исследованиях.

Методы решения задачи

В качестве основы для нахождения решающих правил, разделяющих рассматриваемые 2 класса и позволяющих осуществлять прогноз рецидива, на данном этапе исследований был использован метод и программный комплекс «Фрагмент-потенциал», разработанный в Институте проблем передачи информации им. А.А. Харкевича РАН, в связи с решением задач диагностики и прогнозирования в медицине и обладающий широким диапазоном свойств и возможностей для решения задач в различных предметных областях.

Метод «Фрагмент-потенциал» основан на поиске подмножеств показателей (признаков) и правил попарно разделяющих классы (фрагмент) и последующем голосовании таких правил для принятия решения в «пространстве голосов» (потенциал) [20—23].

Идея алгоритма «Фрагмент» заключается в следующем: по учебной выборке, в которой известна принадлежность объектов к классам, в пространстве описаний для каждого показателя может быть найдена точка (узел), по одну из сторон от которой оказываются объекты какого-либо одного из классов, а по другую — все остальное (практически, однако, алгоритм допускает и не совсем «чистые» разделения). Условие, при котором объекты одного класса располагаются по одну сторону от выбранной точки, запоминается как условие выполнения правила узла. При обучении объекты, удовлетворяющие условию узла, отделяются и далее при построении данного правила не рассматриваются. Процесс продолжается до тех пор, пока не будут разделены все объекты обоих классов обучающей выборки, или дальнейшее разделение не будет возможно, или не будет выполнено ограничение по количеству показателей, входящих в правило. Алгоритмом допускается попадание в отделяемые объекты объектов другого класса, их доля определяется задаваемым параметром. Выбор показателей осуществляется по «качеству» разделения — при определении узла выбирается показатель, лучший по некоторому заданному критерию разделения классов. Простейший показатель качества — доля отделяемых объектов. Последовательное применение разделений объектов классов приводит к созданию кусочно-линейного решающего правила (разделителя). Таким образом, алгоритм «Фрагмент» позволяет находить маломерные (обычно их размерность до 5) просто интерпретируемые кусочно-линейные решающие правила на основании известных характеристик объектов.

Каждое решающее правило — это последовательность узлов, каждый из которых содержит имя и значение некоторого признака из выборки данных, условие «⩽» или «>» для вычисления принадлежности объекта к классу, за который голосует данный узел, номер этого класса (неравенство в узле определяет с какой стороны от значения признака лежат убираемые объекты). В алгоритме настраиваются следующие параметры: минимальное и максимальное количество узлов в правиле, количество разрешенных отделяемых объектов альтернативного класса, количество генерируемых правил.

Пример структуры и свойств правила представлены в виде **табл. 2**. В выборке 101 пациент, из них объектов 1-го класса в выборке (пациенты без рецидива) — 54, объектов 2-го класса в выборке (пациенты с рецидивом) — 47.

Таблица 2. **Пример структуры и свойств правила**
Table 2. **Example of rule structure and properties**

Узел	Действие	Окл	$N_{окл}$	$N_{альт}$	$N_{ост_окл}$	$N_{ост_альт}$
1	nodle >6	2	25	3	22	51
2	size ⩽0,3	2	1	0	21	51
3	tumor ⩽2	2	1	0	20	51
4	cont_nodes ⩽1	1	35	4	16	16
5	n_mtst ⩽2	2	10	2	6	14
6	size ⩽4	1	14	4	2	0

Примечание. Окл — номер отделяемого класса (1∨2); $N_{окл}$ — количество убираемых объектов отделяемого класса; $N_{альт}$ — количество убираемых объектов альтернативного класса; $N_{ост_окл}$ — текущее количество оставшихся объектов в отделяемом классе; $N_{ост_альт}$ — текущее количество оставшихся объектов в альтернативном классе.

В приводимом эксперименте программа построила 11 правил, при этом каждое правило может включать не более 7 показателей (задается параметром). Во все правила программа отобрала следующие 13 показателей: age, size, mets, uzi_nodes, cont_nodes, eho_nodes, qty_nodes, tumor, nodle, n_mtst, ca_bl, betesda, tirads (**см. приложение 1**).

Рассмотрим генерацию множества решающих правил («голосов»). Для поиска множества «голосов» применяется поиск с вычеркиванием. Для этого перед поиском следующего решающего правила исключаются признаки, с которых начинался синтез предшествующих правил. Таким образом, хотя каждое решающее правило маломерно, в получаемый набор «голосов» включается значительная часть показателей исходного пространства и тем самым реализуется избыточность, обеспечивающая надежность результатов.

Отметим, что исключение части признаков из рассмотрения в силу градиентного характера поиска алгоритмом «Фрагмент» решающего правила, не сильно влияет на возможности последующего синтеза. Нормой при этом является полная неповторимость последующих решающих правил, что особенно хорошо заметно при поиске таким способом первых решающих правил, когда свободных признаков еще много.

Кроме того, у алгоритма «Фрагмент» имеются различные возможности, основанные на вариациях основных параметров программы, определяющих конкретный вид оценки качества, приоритеты классов, расположение узлов, допустимую долю пропуска «чужих» при отделении объектов и т. д. Для нахождения решающих правил были опробованы различные варианты значений параметров. В приводимом эксперименте из последующего применения исключались показатели, выбираемые первыми на предшествующих шагах.

В исследовании применен метод классификации, когда решение принимается в результате голосования по некоторому множеству решающих правил. Метод основан на следующей гипотезе: если объект обладает большей частью закономерностей, которым удовлетворяет весь класс, то он удовлетворяет этому классу. Множество закономерностей может быть найдено лишь в том случае, если пространство показателей обладает избыточностью. Алгоритм голосования широко используется в задачах медицинской диагностики и прогнозирования.

Ниже приведены результаты экзамена для 3 пациентов, указаны показатели в правилах, по которым программой принимается решение о принадлежности к одному из классов, а также указан класс. Алгоритм допускает вхождение показателей в несколько правил. Приведем примеры голосования решающих правил, относящих пациента к одному из классов или отказу от классификации («не известно»).

Программа сгенерировала 11 решающих правил с максимальным числом узлов (показателей) равным 7, однако для каждого прецедента число сработавших правил может быть разным. В приводимом эксперименте решение принимается в пользу одного из классов при условии, если разница между голосами, отданными за классы, не менее 5 (порог может меняться), в противном случае, решение о принадлежности к одному из классов не принимается. Каждое правило представлено своими узлами», приведены узлы правил до момента принятия решения по данному пациенту.

Пример 1. Пациент Id=1, истинный класс — 2, экзамен (голосование) класс — 2. Все правила отнесли пациента ко 2-му классу — прогнозируется рецидив. В решение вошли 9 правил, все правила проголосовали за 2-й класс:

mets>0+tumor⩽10+tirads⩽7+age>18++age⩽68+size⩽3 2 класс
qty_nodes>0+tumor⩽10+tirads⩽7+age>18++age⩽68+size⩽3 2
uzi_nodes>1+tumor⩽10+age⩽68+age>18++tirads⩽7+size⩽3 2
eho_nodes>1+tumor⩽10+age⩽68+age>18++tirads⩽7+size⩽3 2
nodle>6 2
size>0.3+cont_nodes>1+age⩽68+age>18++tirads⩽7+n_mtst⩽2 2
cont_nodes>1+age⩽68+age>18+tirads⩽7+n_mtst⩽2 2
age⩽68+tirads⩽7+ca_bl>5 2
tirads⩽7+ca_bl>0+ca_bl⩽7+ca_bl>5 2

Кодированные значения вошедших в правила показателей пациента:

id	num	age	size	mets	uzi_ nodes	cont_ nodes	eho_ nodes	qty_ nodes	tumor	node	n_ mtst	ca_ bl	betesda	tirads	reced
1	65780/ A21	54	2,0	10	2	2	4	9	7	7	1	6	3	7	2

Решающее правило можно условно записать в виде конъюнкции четырех неравенств соответствующих узлов, относящих прецедент к одному из классов (size — размер опухоли, nodle — N:TNM, n_mtst — метастазы M:TNM, ca_bl — кальцитонин): size⩽3 ∧ nodle>6 ∧ n_mtst⩽2 ∧ ca_bl>5.

Пример 2. Пациент Id=11, истинный класс — 1, экзамен (голосование) класс — 1. Все правила отнес-

ли пациента 1-му классу — рецидив не прогнозируется. В решение вошли 8 правил, все правила проголосовали за 1-й класс:

mets>0+tumor>10	1 класс
qty_nodes>0+tumor>10	1
uzi_nodes>1+tumor>10	1
eho_nodes>1+tumor>10	1
nodle⩽6+size>0.3+tumor>2+cont_nodes⩽1	1
tumor>3+size>0.3+cont_nodes⩽1	1
size>0.3+cont_nodes⩽1	1
cont_nodes⩽1	1

Значения вошедших в правила показателей пациента:

id	num	age	size	mets	uzi_ nodes	cont_ nodes	eho_ nodes	qty_ nodes	tumor	node	n_ mtst	ca_ bl	betesda	tirads	reced
11	12116/ C20	43	4,0	3	2	1	4	4	12	5	1	2	3	4	1

Решающее правило можно условно записать в виде конъюнкции значений двух соответствующих узлов, относящих прецедент к одному из классов (tumor — T:TNM, cont_nodes — контур л/у): tumor>10 ∧ cont_nodes=1.

Пример 3. Пациент Id=19, истинный класс — 1, экзамен (голосование) класс не определен, 4 правила отнесли пациента к 1-му классу, 3 правила — ко 2-му классу. В решение вошли 7 правил:

mets>0+tumor>10	1 класс
qty_nodes>0+tumor>10	1
uzi_nodes>1+tumor>10	1
eho_nodes>1+tumor>10	1
nodle⩽6+size>0.3+tumor>2+cont_nodes>1+n_mtst⩽2	2
size>0.3+cont_nodes>1+age⩽68+age>18+tirads⩽7+n_mtst⩽2	2
cont_nodes>1+age⩽68+age>18+tirads⩽7+n_mtst⩽2	2

Значения вошедших в правила показателей пациента:

id	num	age	size	mets	uzi_ nodes	cont_ nodes	eho_ nodes	qty_ nodes	tumor	node	n_ mtst	ca_bl	betesda	tirads	reced
19	14454/ C20	28	1,0	4	2	2	4	4	13	2	1	4	4	6	1

Решающее правило можно записать в виде конъюнкции двух показателей (tumor — T:TNM, n_mtst — M:TNM): tumor>10 ∧ n_mtst⩽2.

Выполнен экзамен на обучающей выборке и методом скользящего контроля. Результаты экзамена на учебной выборке представлены в **табл. 3**. Класс 1 — нет рецидива, класс 2 — рецидив.

Исходя из результатов в **табл. 3**, определяется порог (Level), используемый для прогноза в рабочем режиме. Для описываемого эксперимента целесообразно значение порога 2 или 3. При Level=2 доля правильно классифицированных рецидивов (вероятность обнаружения рецидива, чувствительность) TPR=True_2/Sum_2=43/47=0,91, доля «пропущенных» случаев рецидива (вероятность пропуска цели) FNR=1-TPR=0,09, доля правильно классифицированных «не рецидивов» (специфичность) TNR=True_1/Sum_1=42/54=0,78, доля (вероятность) ложных тревог (рецидива нет, а прогноз рецидива есть) FPR=1-TNR=0,22. Доля ошибочных классификаций Perr=(False_1+ False_2)/(Sum_1 +Sum_2)=(10+3)/(54+47)=0,13.

Другими словами, чувствительность прогностического алгоритма достигает 91% при специфичности 78%, а вероятность ошибки составляет 13%.

Метод скользящего контроля

Скользящий контроль или кросс-проверка или кросс-валидация (cross-validation, CV) — процедура эмпирического оценивания качества алгоритмов, обучаемых по прецедентам. Фиксируется некоторое множество разбиений (обычно 5—10) исходной выборки на две подвыборки: обучающую и контрольную. Для каждого разбиения выполняется настройка алгоритма по обучающей подвыборке, затем оценивается его средняя ошибка на объектах контрольной подвыборки. Оценкой скользящего контроля называется средняя по всем разбиениям величина ошибки

Таблица 3. **Результаты экзамена на учебной выборке**
Table 3. **Exam on the training sample**

Level	True	Error	True_1	True_2	False_1	False_2	Unm_1	Unm_2	Sum_1	Sum_2
1	86	13	43	43	10	3	1	1	54	47
2	85	13	42	43	10	3	2	1	54	47
3	81	12	42	39	9	3	3	5	54	47
4	79	6	42	37	3	3	9	7	54	47
5	79	5	42	37	3	2	9	8	54	47

Примечание. Level — порог (разница между количеством голосующих правил за каждый класс); True — количество правильно классифицируемых пациентов в обоих классах; Error — количество ошибок в двух классах; True_1 и True_2 — количество правильно классифицируемых пациентов за 1-й и 2-й классы соответственно; False_1 и False_2 — количество ошибок за 1-й и 2-й классы соответственно; Unm_1 и Unm_2 — «не знаю» за 1-й и 2-й классы соответственно; Sum_1 и Sum_2 — количество пациентов в 1-м и 2-м классах соответственно.

на контрольных подвыборках. Если выборка независима, то средняя ошибка скользящего контроля дает несмещенную оценку вероятности ошибки. Скользящий контроль является стандартной методикой тестирования и сравнения алгоритмов классификации, регрессии и прогнозирования.

В наших экспериментах мы использовали 10 разбиений, таким образом в контрольную выборку включались 10 пациентов, случайно отобранных из всей выборки. Для каждой группы проводилось обучение на оставшихся 91 пациенте, подобным путем было выполнено 10 испытаний.

Пример одного испытания — скользящий контроль №3.

Экзаменационная (контрольная) выборка:

num obj	37	43	68	83	57	48	86	20	27	15
true_cl	2	1	2	1	2	1	1	1	2	2
exm_cl	2	2	2	1	2	2	1	1	0	2,

где num obj — номер пациента из полной выборки, true_cl — истинный класс, exm_cl — результат экзамена для каждого пациента.

Результат экзамена:

	ex_1Cl	ex_2C	unkn_cl	sum
1Cl	4	1	0	5
2Cl	0	4	1	5

Ntrue=8 Nerr=1 Nunkn=1,

где ex_1Cl, ex_2Cl — результат экзамена в соответствующих классах, unkn_cl — класс «не знаю», sum — количество пациентов в каждом классе, Ntrue — количество правильных, Nerr — количество неправильных, Nunkn — количество «не знаю».

Агрегированные результаты 10 испытаний методом скользящего контроля представлены в **табл. 4**.

Суммарные показатели по 2-м классам:

Ntrue=80, Nerr=15, Nunkn=5.

Суммарная доля правильно классифицированных рецидивов по 10 испытаниям скользящего контроля (вероятность обнаружения рецидива, чувствительность)

TPR=ex_2Cl_Total/Sum_2Total=42/49=0.86,

доля «пропущенных» случаев рецидива (вероятность пропуска цели) FNR=1-TPR=0,14, суммарная доля правильно классифицированных «не рецидивов» (специфичность)

TNR=ex_1Cl_Total/Sum_1Total=38/51=0,75,

суммарная доля (вероятность) ложных тревог (рецидива нет, а прогноз рецидива есть) FPR=1-TNR=0,25, суммарная доля ошибочных классификаций

Perr=(ex_2Cl_1Total+ex_1Cl_2Total)/(Sum_1Total+Sum_2Total)=(13+2)/(51+49)=0,15.

Таким образом, при экзамене методом скользящего контроля чувствительность прогностического алгоритма достигает 86% при специфичности 75% и вероятности ошибки 15%. Полученные результаты близки к результатам по полной выборке, что подтверждает эффективность используемого метода для прогноза рецидивов.

Заключение

Таким образом, выполненные пилотные эксперименты показали наличие закономерностей в дан-

Таблица 4. **Агрегированные результаты 10 испытаний методом скользящего контроля**
Table 4. **Aggregated results of 10 tests using the sliding control method**

	ex_1Cl	ex_2Cl	unkn_cl	Sum		ex_1Cl	ex_2Cl	unkn_cl	Sum
1Cl	5	2	0	7	2Cl	0	3	0	3
1Cl	3	1	0	4	2Cl	1	5	0	6
1Cl	4	1	0	5	2Cl	0	4	1	5
1Cl	3	0	0	3	2Cl	1	5	1	7
1Cl	4	1	0	5	2Cl	0	4	1	5
1Cl	3	3	0	6	2Cl	0	4	0	4
1Cl	3	2	0	5	2Cl	0	4	1	5
1Cl	4	2	0	6	2Cl	0	4	0	4
1Cl	4	1	0	5	2Cl	0	5	0	5
1Cl	5	0	0	5	2Cl	0	4	1	5
1Total	38	13	0	51	2Total	2	42	5	49

Суммарные результаты по всем группам

	ex_1Cl	ex_2Cl	unkn_cl	sum
1Cl	38	13	0	51
2Cl	2	42	5	49

ных, позволяющих решать задачу прогноза рецидивов по показателям дооперационного обследования, и перспективность применяемого метода. Продолжение исследований предполагает увеличение выборки минимум до 200 пациентов, привлечение неформализованных врачебных знаний для формирования решений, применение других методов классификации.

Дальнейшее изучение проблемы может помочь врачам в принятии решений, а также внести значительные коррективы в тактику ведения пациентов с ПРЩЖ.

Авторы заявляют об отсутствии конфликта интересов.
The authors declare no conflicts of interest.

ЛИТЕРАТУРА/REFERENCES

1. *Персонализированная эндокринология в клинических примерах*. Под ред. И.И. Дедова. М.: «ГЭОСТАР-Медиа»; 2019.
Personalized endocrinology in clinical examples. II Dedov, ed. M.: GEOSTAR-Media Publishing Group; 2019. (In Russ.).
2. Каприн А.Д. *Состояние онкологической помощи населению России в 2021 году*. М.: МНИОИ им. П.А. Герцена – филиал ФГБУ «НМИЦ радиологии» Минздрава России; 2022.
Kaprin AD. *The state of cancer care for the population of Russia in 2021*. M.: MNIOI im. P.A. Herzen - branch of the Federal State Budgetary Institution «National Medical Research Center of Radiology» of the Ministry of Health of Russia; 2022. (In Russ.).
3. Lloyd R, Osamura R, Klöppel G, Rosai J. *WHO classification of tumours of endocrine organs. 4*. Lyon: International Agency for Research on Cancer; 2017.
4. Clark OH. Thyroid cancer and lymph node metastases. *Journal of Surgical Oncology*. 2011;103:615-618.
5. Скибицкая М.В., Кузнецов Н.С. Современные аспекты, влияющие на тактику ведения пациентов с папиллярным раком щитовидной железы. *Хирургия. Журнал им. Н.И. Пирогова*. 2023;(12):89-94.
Skibitskaya MV, Kuznetsov NS. Modern aspects influencing the management of patients with papillary thyroid cancer. *Khirurgiya. Zhurnal im. N.I. Pirogova*. 2023;(12):89-94. (In Russ.).
https://doi.org/10.17116/hirurgia202312189
6. Ahn SH, Kim WS. The Effect of Prophylactic Central Neck Dissection During Hemithyroidectomy on Locoregional Recurrence in Patients With Papillary Thyroid Carcinoma: A Meta-Analysis. *Clinical and Experimental Otorhinolaryngology*. 2020;13(2):194-202.
https://doi.org/10.21053/ceo.2019.01634
7. Coca-Pelaz A, Shah JP, Hernandez-Prera JC, Ghossein RA, Rodrigo JP, Hartl DM, Olsen KD, Shaha AR, Zafereo M, Suarez C, Nixon IJ, Randolph GW, Mäkitie AA, Kowalski LP, Vander Poorten V, Sanabria A, Guntinas-Lichius O, Simo R, Zbären P, Angelos P, Khafif A, Rinaldo A, Ferlito A. Papillary Thyroid Cancer-Aggressive Variants and Impact on Management: A Narrative Review. *Advances in Therapy*. 2020;37(7):3112-3128.
https://doi.org/10.1007/s12325-020-01391-1
8. Clark OH. Thyroid cancer and lymph node metastases. *Journal of Surgical Oncology*. 2011;103(6):615-8.
https://doi.org/10.1002/jso.21804
9. Lam AK, Lo CY, Lam KS. Papillary carcinoma of thyroid: A 30-yr clinicopathological review of the histological variants. *Endocrine Pathology*. 2005;16(4):323-30.
https://doi.org/10.1385/ep:16:4:323
10. Кожанов Л.Г., Сдвижков А.М., Романова Е.С. Ошибки в диагностике и лечении рака щитовидной железы. *Российский онкологический журнал*. 2019;24(1-2):58-59.
Kozhanov LG, Sdvizhkov AM, Romanova ES. Errors in the diagnosis and treatment of thyroid cancer. *Russian Journal of Oncology*. 2019;24(1-2):58-59. (In Russ.).
11. Заридзе Д.Г., Стилиди И.С., Каприн А.Д., Максимович Д.М. Гипердиагностика рака щитовидной железы. *Практическая онкология*. 2020;21(4):312-326.
Zaridze DG, Stilidi IS, Kaprin AD, Maksimovich DM. Overdiagnosis of thyroid cancer. *Practical oncology*. 2020;21(4):312-326. (In Russ.).
https://doi.org/10.31917/2104312
12. Бельцевич Д.Г., Ванушко В.Э., Румянцев П.О., Мельниченко Г.А., Кузнецов Н.С., Абросимов А.Ю., Поляков В.Г., Мудунов А.М., Подвязников С.О., Романов И.С., Поляков А.П., Слепцов И.В., Черников Р.А., Воробьев С.Л., Фадеев В.В. Российские клинические рекомендации по диагностике и лечению высокодифференцированного рака щитовидной железы у взрослых, 2017 год. *Эндокринная хирургия*. 2017;11(1):6-27.
Beltsevich DG, Vanushko VE, Rumyantsev PO, Melnichenko GA, Kuznetsov NS, Abrosimov AYu, Polyakov VG, Mudunov AM, Podvyaznikov SO, Romanov IS, Polyakov AP, Sleptsov IV, Chernikov RA, Vorobyov SL, Fadeev VV. Russian clinical guidelines for the diagnosis and treatment of highly differentiated thyroid cancer in adults, 2017. *Endocrine surgery*. 2017;11(1):6-27.
https://doi.org/10.14341/serg201716-27]
13. Ahn SH, Kim WS. The Effect of Prophylactic Central Neck Dissection During Hemithyroidectomy on Locoregional Recurrence in Patients With Papillary Thyroid Carcinoma: A Meta-Analysis. *Clinical and Experimental Otorhinolaryngology*. 2020;13(2):194-202.
https://doi.org/10.21053/ceo.2019.01634
14. Eggermont AM. Reducing the Need for Sentinel Node Procedures by Ultrasound Examination of Regional Lymph Nodes. *Annals of Surgical Oncology*. 2005;12:3-5.
15. Esen G. Ultrasound of superficial lymph nodes. *European Journal of Radiology*. 2006;58(3):345-359.
16. Cui XW, Hocke M, Jenssen C, Ignee A, Klein S, Schreiber-Dietrich D, Dietrich CF. Conventional ultrasound for lymph node evaluation, update 2013. *Zeitschrift für Gastroenterologie*. 2014;52(2):212-21.
https://doi.org/10.1055/s-0033-1356153
17. Ryu KH, Lee KH, Ryu J, Baek HJ, Kim SJ, Jung HK, Kim SM. Cervical Lymph Node Imaging Reporting and Data System for Ultrasound of Cervical Lymphadenopathy: A Pilot Study. *American Journal of Roentgenology*. 2016;206(6):1286-91.
https://doi.org/10.2214/AJR.15.15381
18. Щеголев А.И., Туманова У.Н., Кармазановский Г.Г. Восьмое издание Международной TNM-классификации злокачественных опухолей печени. *Медицинская визуализация*. 2017;(6):41-48.
Shchegolev AI, Tumanova UN, Karmazanovsky GG. Eighth Edition of the International TNM Classification of Malignant Tumors of the Liver. *Medical Visualization*. 2017;(6):41-48. (In Russ.).
https://doi.org/10.24835/1607-0763-2017-6-41-48
19. Фисенко Е.П., Сыч Ю.П., Ветшева Н.Н. К вопросу о классификации TI-RADS и стратификации признаков рака щитовидной железы по данным ультразвукового исследования. *Медицинская визуализация*. 2017;(5):29-38.
Fisenko EP, Sich JP, Vetsheva NN. On the Classification of TI-RADS and Stratification of Signs of Thyroid Cancer According to Ultrasound Data. *Medical Visualization*. 2017;(5):29-38. (In Russ.)
https://doi.org/10.24835/1607-0763-2017-5-29-38
20. Левит В.Е., Переверзев-Орлов В.С. *Структура и поле данных при распознавании образов*. М.: Наука; 1984.
Levit VE, Pereverzev-Orlov VS. *Structure and data field in pattern recognition*. M: Science; 1984. (In Russ.).
21. Вайнштok А.П., Переверзев-Орлов В.С. Программный комплекс «Распознавание образов с одновременным поиском подмножеств признаков и правил попарно разделяющих классы». *ГФАП СССР. Алгоритмы и программы. Информационный бюллетень*. 1985;4(67).

Vainshtok AP, Pereverzev-Orlov VS. Software package «Pattern recognition with simultaneous search for subsets of features and rules separating classes in pairs». *SFAP USSR. Algorithms and programs. Information Bulletin*. 1985;4(67). (In Russ.).

22. Vashenko E, Vitushko M, Pereverzev-Orlov V. Potentials of Learning on the Basis of Partner System. *Pattern Recognition and Image Analysis*. 2004;14(1):84-91.

23. Ващенко Е.А., Витушко М.А., Переверзев-Орлов В.С., Стенина И.И. Советчик врача: технологии и возможности. *Электронный научный журнал «Информационные Процессы»*. 2009;9(1):18-24. Vashenko EA, Vitushko MA, Pereverzev-Orlov VS, Stenina II. Doctor's advisor: technologies and capabilities. *Electronic scientific journal Information Processes*. 2009;9(1):18-24. (In Russ.).

Поступила 29.03.2024
Received 29.03.2024
Принята к печати 22.04.2024
Accepted 22.04.2024

Приложение/Application 1
Полный список введенных в ЭВМ показателей, их идентификаторы и кодировка

1. Номер карты: num.
2. Пол [женский — 1; мужской — 2]: pol.
3. Возраст (18—110): age (1—110).
4. 1-я диагностика, год: year 1.
5. Дата операции: oper_dat.
6. УЗИ опухоли:
   — размер опухоли, см: size;
   — контур — неровность\нечеткость [нет — 1; да — 2]: cont;
   — кальцинаты [нет — 1; да — 2]: Ca;
   — эхоструктура [изоэхогенная — 1; гиперэхогенная — 2; гипоэхогенная — 3]: eho.
7. Метастазы, кол-во: mets.
8. УЗИ лимфатических узлов, патология [НЕТ — 1; ДА — 2]: uzi_nodes:
   — контур — неровность\нечеткость [нет — 1; да — 2]: cont_nodes;
   — кальцинаты [нет — 1; да — 2]: Ca_nodes;
   — эхоструктура [изоэхогенная — 2; гиперэхогенная — 3; гипоэхогенная — 4]: eho_nodes;
   — количество метостатических узлов: qty_nodes.
9. TNM:
   T [T –1; Tx –2; T0 –3; T1 –4; T1a –5; T1b –6; T2 –7; T3 –8; T3a –9;
   T3b –10; T4 –11; T4a –12; T4b –13]: tumor
   N [Nx –1; N0 –2; N0a –3; N0b –4; N1 –5; N1a –6; N1b –7]: nodle
   M (metastases) [M0 –1; M1 –2; MX –3]: n_mtst.
10. Кальцитонин, кровь пг/мл: ca_bl.
11. BETESDA — диагностическая категория, риск %. [Недиагностический или неудовлетворительный пунктат (1—4) — 1; доброкачественный (0—3) — 2; атипичная или фолликулярные изменения неопределенного значения (5—15) — 3; фолликулярная или подозрение на фолликулярную неоплазию (15—30) — 4; подозрение на рак (60—75) — 5; Рак (97—99) — 6]: betesda.
12. TIRADS [TIRADS 1-1; TIRADS 2-2; TIRADS 3-3; TIRADS 4a-4; TIRADS 4b-5; TIRADS 4c-6; TIRADS 5-7; TIRADS 6-8]: tirads.
13. Послеоперационные осложнения: [НЕТ — 1; ДА — 2]: post_oper.
14. Осложнение [кровотечение — 2; парез возвратного нерва — 3; гипокальцемия — 4]: postoper.
15. Сосудистая инвазия (морфология) [НЕТ — 1; ДА — 2]: sosud_inv.
16. Количество пораженных лимфатических узлов (морфология): qty_morf.
17. Геми+ипсилатеральная лимфоаденоктомия [НЕТ — 1; ДА — 2]: ipsi.
18. Рецидив [НЕТ — 1; ДА — 2]: reced.
19. Дата рецидива: data_rec.