

**ТРОХИНА И. Е., ЗАМЫШЛЯЕВ П. С., ЯМАШКИНА В. В.,**

**МОРОЗКИНА А. А., ФЕДИН А. А.**

**РАЗРАБОТКА ПРОГНОСТИЧЕСКОЙ МОДЕЛИ  
ЯЗВЕННОЙ БОЛЕЗНИ ДВЕНАДЦАТИПЕРСТНОЙ КИШКИ  
С ПРИМЕНЕНИЕМ КОМПЬЮТЕРНОГО АНАЛИЗА ДАННЫХ**

**Аннотация.** Среди заболеваний органов пищеварения язвенная болезнь двенадцатиперстной кишки продолжает оставаться одной из наиболее частых причин обращаемости людей за медицинской помощью. Проблема прогноза возможности возникновения и особенностей течения заболевания всегда была и остается важнейшим разделом в деятельности врачей и предметом интенсивной активности ученых в поиске решения задачи. В статье представили этапность и результат разработки прогностической модели язвенной болезни двенадцатиперстной кишки с применением компьютерного анализа данных.

**Ключевые слова:** искусственный интеллект, язвенная болезнь, прогноз заболеваний.

**TROKHINA I.E., ZAMYSHLYAEV P.S., YAMASHKINA V.V.,**

**MOROZKINA A.A., FEDIN A.A.**

**DEVELOPMENT OF A PROGNOSTIC AND DIAGNOSTIC MODEL  
OF DUODENAL ULCER USING COMPUTER DATA ANALYSIS**

**Abstract.** Duodenal ulcer among diseases of the digestive system continues to be one of the most common reasons for people seeking medical help. The problem of predicting the possibility of the occurrence and characteristics of the course of a disease has always been and remains the most important section in the work of doctors and the subject of intensive activity of scientists in finding a solution to the problem. The article presented the stages and results of developing a prognostic model of duodenal ulcer using computer data analysis.

**Keywords:** artificial intelligence, duodenal ulcer disease, disease prognosis.

**Введение.** Искусственный интеллект – это область информатики, занимающаяся разработкой систем, которые способны выполнять манипуляции, традиционно связанные с человеческим ресурсом. Сегодня все чаще ведутся дискуссии о преимуществах и недостатках использования искусственного интеллекта в различных профессиональных сферах, в том числе и в медицине [1].

Умение мыслить и действовать нестандартно в сложных ситуациях играет главную роль в правильной постановке диагноза. Это зачастую связано с тем, что знания и опыт врача имеют в значительной степени нестатистический характер, который является результатом

взаимодействия теоретических, практических знаний и собственных клинических наблюдений. Именно таким особенностям современной медицины полностью отвечают интеллектуальные системы (ИС), которые основаны на анализе данных и знаниях высоко профессиональных специалистов-экспертов, а не на информации о пациентах (отсюда их первое название – экспертные системы) [2].

Язвенная болезнь (ЯБ) двенадцатиперстной кишки (ДПК), в современном обществе, остается одной из частых причин обращаемости людей за медицинской помощью несмотря на постоянное усовершенствование различных методов обследования и лечения больных с данной патологией [3]. Неоднократно доказано, что большое количество язв ДПК имеет хроническое и рецидивирующее течение, поэтому прогнозирование течения ЯБ приобретает первостепенное значение в определении лечебной тактики и предупреждении осложнений (перфорация, кровотечение, рубцовые стенозы привратника, грубая деформация выходного отдела желудка и ДПК) [4].

Много остается неясного в различных механизмах рецидивов и увеличения риска и частоты осложнений заболевания после, казалось бы, успешной терапии. Кроме того, если говорить о впервые выявленных случаях ЯБ, то прогноз их течения является наиболее актуальным на данном этапе [5]. Исходя из этого, существует большая необходимость модернизации и более полного и активного использования различных компьютерных диагностических технологий, которые необходимо направить на прогнозирование течения ЯБ [6].

Известно, что ранее уже разрабатывались специализированные информационные комплексы, которые позволяют оценить факторы риска с использованием традиционных биостатистических методов, но в полной мере решить проблему разработки персонализированной профилактической стратегии они не могут в силу свойственных им фундаментальных математических ограничений. В то же время ИС продемонстрировали большие возможности в анализе сложных данных [7], однако их потенциал в рассматриваемой области медицины только предстоит изучить.

**Цель работы:** разработать прогностическую модель язвенной болезни двенадцатиперстной кишки с применением компьютерного анализа данных.

**Материалы и методы.** Анализировали данные пациентов с язвенной болезнью и данные здоровых лиц, собранных ранее в рамках диссертационного исследования канд. мед. наук И. Е. Трохиной.

Диагноз определялся экспертно на основании анамнестических, клинических данных, результатов лабораторных и инструментальных методов исследования – эзофагогастродуоденоскопии (ЭФГДС), исследования на инфицированность *Helicobacter*

*pylori*. ЭФГДС проводили на основе общепринятой методики эндоскопами Olympus, Exera (CIF160) в утренние часы, натощак. Анализировались данные лабораторных методов исследования (общего и биохимического анализов крови) при поступлении, на 7-е, 14-е, 30-е сутки течения заболевания. Кровь для исследования брали при первичном обращении до начала лечения и в течение времени при проведении терапии при получении информированного добровольного согласия обследуемых лиц. Использовались данные 27 здорового добровольца (группа 1), 28 больных с ЯБ ДПК (группа 2) и 28 больных с ЯБ ДПК, страдающих хронической ишемией головного мозга (группа 3).

Собранные обезличенные данные анализировались в среде JupyterLab с установленным ядром-интерпретатором языка Python версии 3.8 на компьютере (рабочей станции) с операционной системой Linux Ubuntu 22.04.3 LTS, применялись стандартные библиотеки для анализа данных.

Проводили сначала анализ объединенных данных групп 1 и 2 (то есть, здоровых лиц и больных ЯБ без хронической ишемии мозга), а затем объединенных данных групп 2 и 3 (то есть, только по отношению к данным больных ЯБ, но опционально дополнительно страдающих хронической ишемией головного мозга).

**Результаты.** Для построения прогностической модели данные были полуавтоматически (для исключения ошибок ручного редактирования) предобработаны: были удалены столбцы, по которым данные были однородны по двум группам (например, столбец с результатами теста на *Helicobacter pylori*, который у всех пациентов в выборке был положительный), проводился контроль пропусков значений (пропущенных значений не было выявлено). Собственно, исследовательский анализ данных после предобработки включал проверку характера данных, корреляционный анализ и применение метода главных компонент.

Проверка характера данных заключалась в оценке характера распределения записанных показателей исследуемых лиц и вычислении классических статистических показателей: среднего значения – mean, стандартного отклонения – std, минимального и максимального значений – min и max, значений, соответствующих 25, 50, 75 перцентилю распределения – 25%, 50%, 75% (рис. 1).

Интересная находка была обнаружена при визуализации данных и заключалась в том, что пациенты на основании двухпиковости распределения размеров язвы при поступлении и распределения интенсивности уменьшения язвы к окончанию второй недели лечения могут быть четко разделены на непересекающиеся фенотипы по следующим критериям: с язвой на момент поступления до 10 мм и более 10 мм, а также с уменьшением размеров язвы на 14 сутки лечения до 1 мм включительно и более 1 мм.

|       | ЛФХ      | Сфингомиелин | Фосфатидилхолин | Фосфатидилсерин | Фосфатидилинозит |
|-------|----------|--------------|-----------------|-----------------|------------------|
| count | 65.00000 | 65.000000    | 65.000000       | 65.000000       | 65.000000        |
| mean  | 1.12000  | 24.800769    | 47.762769       | 8.191385        | 0.734462         |
| std   | 0.01723  | 0.605218     | 0.673699        | 0.150830        | 0.022983         |
| min   | 1.09000  | 23.810000    | 46.060000       | 7.480000        | 0.640000         |
| 25 %  | 1.11000  | 24.660000    | 47.320000       | 8.120000        | 0.730000         |
| 50 %  | 1.12000  | 24.750000    | 48.030000       | 8.140000        | 0.740000         |
| 75 %  | 1.13000  | 24.800000    | 48.300000       | 8.260000        | 0.740000         |
| max   | 1.20000  | 27.800000    | 48.800000       | 8.610000        | 0.770000         |

Рис. 1. Пример таблицы, полученной при проверке характера данных.

Далее представлены примеры визуализации распределений данных (рис. 2).

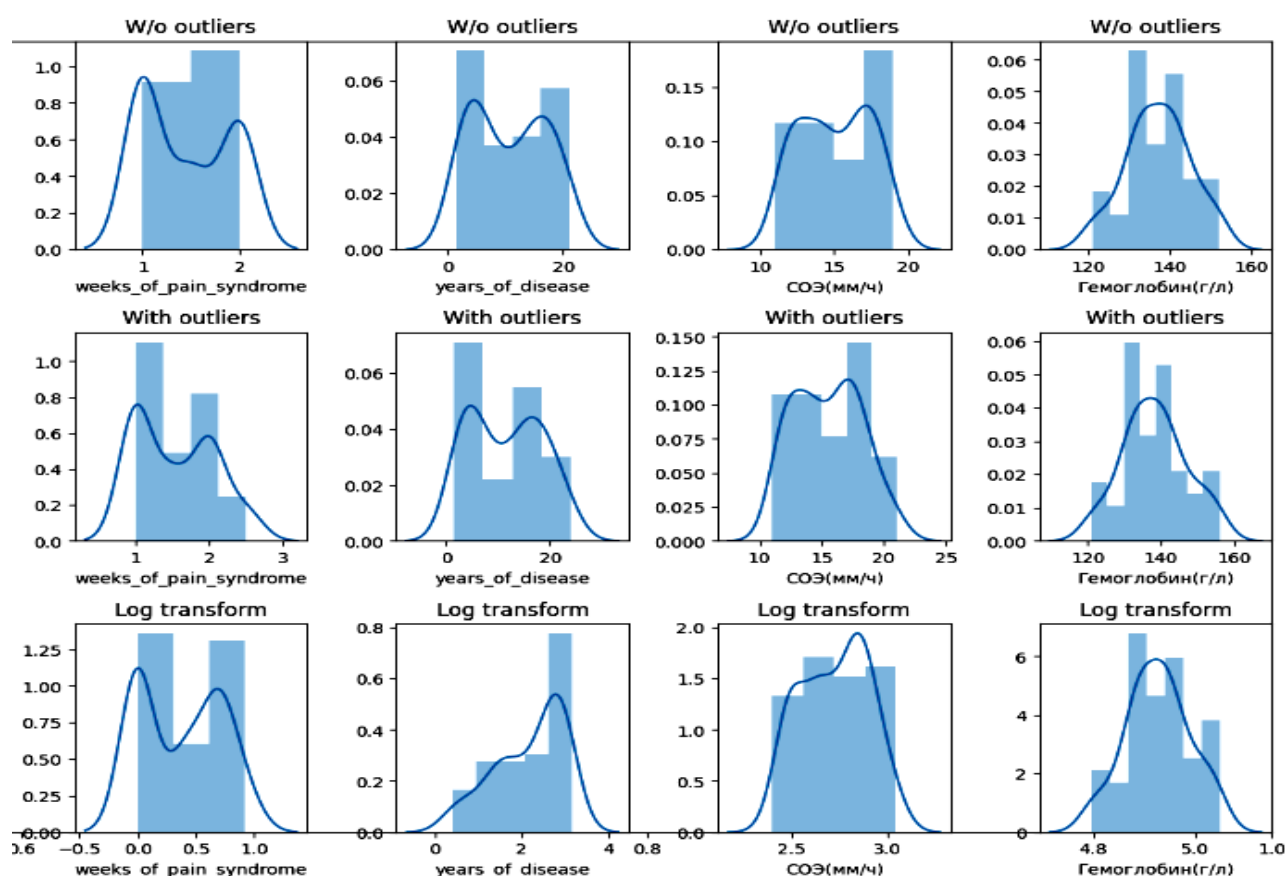


Рис. 2. Примеры гистограмм проверки характера данных.

Примечания: в столбцы сверху вниз распределены по 3 гистограммы

(1-я – распределение без выбросов; 2-я – данные без обработки; 3-я – данные с лог-трансформацией),  
по оси ординат – частота встречаемости конкретного значения (ось абсцисс).

Следующим шагом исследовательского анализа данных стало проведение корреляционного анализа данных пациентов. Были удалены данные по показателям,

коррелирующим с большим числом других показателей (в этом исследовании убирали показатели, имеющие более 3 коррелирующих с ними). Далее приведем примеры таких корреляций в визуальном представлении (рис. 3).

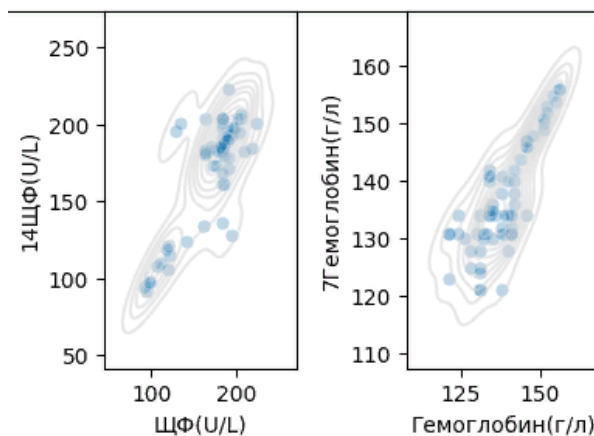


Рис. 3. Примеры визуального представления корреляции между исследованными признаками.

После чистки корреляций, обнаруженных с помощью стандартизации значений факторов и собственно поиска корреляций стандартными методами, проводили анализ с применением метода главных компонент. С помощью этого метода было выяснено, что практически всю вариабельность данных пациентов с ЯБ из выборки нашего исследования возможно соотнести с вариабельностью следующих признаков: длительностью анамнеза курения, а также различиями в показателях роста, массы тела, показателях анализов крови: гемоглобин, билирубин, щелочной фосфатазы (ЩФ) и диацилглицерола (ДАГ).

**Закключение.** Согласно нашим исследованиям, именно такие признаки как длительность анамнеза курения, рост, масса тела, гемоглобин, билирубин, ЩФ, ДАГ – вероятно, наиболее целесообразно использовать на следующих этапах исследований для построения прогностических моделей машинного обучения (искусственного интеллекта), нацеленных на формирование скрининговых заключений относительно ЯБ ДПК, а также построения моделей для выделения фенотипов пациентов.

#### СПИСОК ЛИТЕРАТУРЫ

1. Лазаренко В.А., Антонов А.Е., Marcaruram V.K., Awad K. Опыт нейросетевой диагностики и прогнозирования язвенной болезни по результатам анализа факторов риска // Бюллетень сибирской медицины. – 2018. – № 17 (3). – С. 88–95.
2. Кобринский Б.А. Искусственный интеллект и медицина: Возможности и перспективы систем, основанных на знаниях // Новости искусственного интеллекта. – 2019. – № 4. – С. 44–51.

3. Кобринский Б.А., Хавкин А.И., Волюнец Г.В. Перспективы применения систем искусственного интеллекта в гастроэнтерологии // Экспериментальная и клиническая гастроэнтерология. – 2020. – № 179 (7). – С. 109–117.
4. Миненок В.А. Использование искусственного интеллекта в медицине // М.: Сборник тезисов конференции СВАИ «Вычислительная биология и искусственный интеллект для персонализированной медицины». – С. 75–76.
5. Маев И.В., Андреев Д.Н., Самсонов А.А., Черёмушкина А.С. Язвенная болезнь: современное состояние проблемы // Медицинский совет. – 2022. – № 16 (6). – С. 100–108.
6. Blaser M.J., Atherton J.C. Helicobacter pylori persistence: biology and disease // J. Clin. Invest. – 2020. – Vol. 113, No. 3. – P. 321–333.
7. Грибова В.В., Петряева М.В., Шалфеева Е.А. Облачный сервис поддержки принятия диагностических решений в гастроэнтерологии // Врач и информационные технологии. – 2019. – № 3. – С. 65–71.