



<https://doi.org/10.29001/2073-8552-2025-2706>
УДК 616-039.42-07:004.8

Система искусственного интеллекта для диагностики редких заболеваний: принципы построения и результаты клинической апробации

Б.А. Кобринский^{1, 2}, Н.А. Благосклонов^{1, 2}

¹ Федеральный исследовательский центр «Информатика и управление» Российской академии наук (ФИЦ ИУ РАН), 119333, Российская Федерация, Москва, ул. Вавилова, 44, корп. 2

² Российский национальный исследовательский медицинский университет имени Н.И. Пирогова Министерства здравоохранения Российской Федерации (Пироговский Университет) (РНИМУ им. Н.И. Пирогова (Пироговский Университет)), 117513, Российская Федерация, Москва, ул. Островитянова, 1, стр. 6.

Аннотация

Введение. Дифференциальная диагностика редких болезней на долабораторном этапе обследования пациентов является сложной задачей не только для педиатров, но и для врачей-генетиков. Это обусловлено такими факторами, как недостаток у многих врачей личного опыта в наблюдении пациентов с редкими болезнями, вариативность и нечеткость клинических проявлений отдельных признаков. Интеллектуальные системы поддержки принятия решений обеспечивают выдвижение и обоснование диагностических гипотез.

Цель исследования: рассмотреть архитектуру экспертной системы поддержки принятия решений для дифференциальной диагностики лизосомных болезней накопления (ЛБН) на долабораторном этапе и представить результаты клинической апробации.

Материал и методы. Материалом исследования являлись 30 клинических форм ЛБН, которые были описаны 35 признаками на основе трех взаимно дополняющих экспертных оценок – коэффициента модальности (важности для диагностики), факторов уверенности манифестации и выраженности. Знания были получены из литературных источников, а затем уточнены экспертами, указавшими свою уверенность в каждом значении признака в каждом из 4 возрастных диапазонов (до 1 года, от 1 года до 3 лет включительно, от 4 до 6 лет включительно, 7 лет и старше). Эта информация послужила основой для формирования базы знаний интеллектуальной системы. Материалом для клинической апробации были деперсонифицированные выписки из электронных медицинских карт детей с ЛБН (мукополисахаридозы (МПС), муколипидозы (МЛП), ганглиозидозы (ГЗ), которые известны широкими перекрывающимися фенотипическими спектрами): 54 – для валидации (из одной медицинской организации), 38 – для верификации (из трех медицинских организаций трех регионов Российской Федерации). Для построения интеллектуальной системы были использованы методы инженерии знаний (для извлечения и структуризации знаний), матричный подход (для представления правил базы знаний), информационное и программное обеспечение для реализации архитектуры интеллектуальной системы.

Результаты. Модифицирована экспертная система поддержки принятия решений по дифференциальной диагностике орфанных наследственных заболеваний ГенДиЭС. В новой версии система реализована в виде web-приложения. База знаний содержит 12 600 оценок, характеризующих меры доверия экспертов для 35 признаков 30 клинических форм ЛБН по 4 возрастным периодам. Алгоритм принятия решений обеспечивает оценку сходства клинической картины пациента с экспертными описаниями отдельных ЛБН. Точность для дифференциально-диагностического ряда из пяти гипотез составила 0,87 (95% ДИ [0,75; 0,95]) и 0,90 (95% ДИ [0,75; 0,97]) для валидации и верификации соответственно.

Заключение. Интеллектуальная система ГенДиЭС показала высокую точность на долабораторном этапе диагностики, сопоставимую с мировыми аналогами. Реализация системы в виде web-приложения в свободном доступе предоставляет врачам возможность обращения с любого устройства, подключенного к сети Интернет.

Ключевые слова: система искусственного интеллекта; экспертная система; ГенДиЭС; система поддержки врачебных решений; дифференциальная диагностика; наследственные болезни; лизосомные болезни накопления; валидация; верификация.

Финансирование: исследование выполнено в рамках государственного задания «Системы искусственного интеллекта, извлечения знаний и анализа текстов», государственный регистрационный номер: 124040100006-4, Шифр: FFNG-2024-0001.

✉ Кобринский Борис Аркадьевич, e-mail: kba_05@mail.ru.

Соответствие принципам этики: исследование проведено на обезличенных ретроспективных данных пациентов.

Для цитирования: Кобринский Б.А., Благодосклонов Н.А. Система искусственного интеллекта для диагностики редких заболеваний: принципы построения и результаты клинической апробации. *Сибирский журнал клинической и экспериментальной медицины*. 2025;40(2):218–225 <https://doi.org/10.29001/2073-8552-2025-2706>

Artificial Intelligence System for Diagnosing Rare Diseases: Design Principles and Clinical Validation Results

Boris A. Kobrinskii^{1,2}, Nikolay A. Blagosklonov^{1,2}

¹ Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44, build. 2, Vavilova str., Moscow, 119333, Russian Federation

² Pirogov Russian National Research Medical University, 1, Bldg 6, Ostrovityanova str., Moscow, 117513, Russian Federation

Abstract

Introduction. Differential diagnostics of rare diseases at the pre-laboratory stage of patient examination is a significant challenge not only for pediatricians but also for geneticists. This fact is caused by limited physician experience in managing rare diseases, variability in symptom presentation, and ambiguity of clinical signs. AI-driven clinical decision support systems (CDSS) enable the generation and validation of diagnostic hypotheses.

Aim: To assess the architecture of the GenDiES (Genetic Diagnostic Expert System) CDSS for differential diagnosis of lysosomal storage diseases (LSDs) at the pre-laboratory stage and to present results from its clinical validation.

Material and Methods. The study included 30 clinical forms of LSDs, described using 35 clinical features based on three complementary expert-derived metrics: modality coefficient (diagnostic importance), manifestation certainty factor, and degree of expression certainty factor. Knowledge was extracted from literature and refined through expert input, where experts assigned their confidence to each feature across four age groups: ≤ 1 year, 1–3 years, 4–6 years, ≥ 7 years. This structured data formed the system’s knowledge base. Clinical aprobatation utilized de-identified electronic health records (EHR) of pediatric LSD patients (mucopolysaccharidoses, mucolipidoses, gangliosidoses-conditions with broad overlapping phenotypic spectra). Validation cohort: 54 EHR extracts from a single Russian medical institution. Verification cohort: 38 EHR extracts from three institutions across different Russian regions. The system was built using knowledge engineering methods (for knowledge extraction and structuring), a matrix-based framework (to organize rules), and custom software for implementation.

Results. The updated GenDiES CDSS for differential diagnosis of rare hereditary diseases was deployed as a web application. Its knowledge base contains 12,600 expert confidence assessments for 35 clinical features across 30 LSD subtypes, categorized by age. A similarity-based algorithm compares patient profiles to expert-defined disease patterns. Accuracy for generating a differential diagnosis shortlist (top five hypotheses) reached 0.87 (95% CI [0.75; 0.95]) during validation and 0.90 (95% CI [0.75; 0.97]) during verification.

Conclusion. The GenDiES system demonstrated high diagnostic accuracy at the pre-laboratory stage, comparable to—and in some cases exceeding—the performance of limited existing international counterparts. Its web-based implementation ensures accessibility for physicians via any internet-connected device.

Keywords: artificial intelligence system; expert system; GenDiES; clinical decision support system; differential diagnosis; hereditary diseases; lysosomal storage diseases; validation; verification.

Funding: the study was carried out within the framework of the State assignment “Artificial Intelligence Systems, Knowledge Extraction, and Text Analysis” (State Registration No. 124040100006-4; Code: FFNG-2024-0001).

Compliance with ethical standards: the study used de-identified retrospective patient data.

For citation: Kobrinskii B.A., Blagosklonov N.A. Artificial Intelligence System for Diagnosing Rare Diseases: Design Principles and Clinical Validation Results. *Siberian Journal of Clinical and Experimental Medicine*. 2025;40(2):218–225 <https://doi.org/10.29001/2073-8552-2025-2706>

Введение

Всемирная организация здравоохранения предложила рассматривать частоту менее 10 на 10 000 человек для орфанных заболеваний. В Европейском Союзе (ЕС) это определение обозначает частоту менее чем 5 на 10 000 человек (или 1 на 2 000). Законом США (Rare Disease Act) установлено, что редкое заболевание – это заболевание, которое поражает менее 200 000 человек, в Японии – менее 50 000, в Австралии – менее 2 000 человек [1].

Федеральный закон Российской Федерации № 323-ФЗ от 21 ноября 2011 г. «Об основах охраны здоровья граждан Российской Федерации» определяет для редких болезней распространенность не более чем 1 случай на 10 000 населения. Встречаются еще более редкие клинические формы, например, с частотой 1 : 200 000 [2]. Одной из групп орфанных болезней являются наследственные лизосомные болезни накопления (ЛБН) [3], которые включают мукополисахаридозы (МПС), ганглиозидозы (ГЗ) и муколипидозы (МЛП) и встречаются с частотой от 1 : 40 000 до 1 : 1 000 000 и реже [4]. Всего в данную группу входят около 60 клинических форм [5].

Диагностика редких заболеваний – многоэтапный процесс, в котором задействованы разные специалисты, начиная с первого месяца жизни ребенка. Это обусловлено тем, что несмотря на относительную редкость отдельных нозологических форм, более 50 различных ЛБН проявляются в неонатальный период [6]. Сначала врач-педиатр должен заподозрить наличие генетически обусловленного заболевания по клиническим проявлениям. Затем пациент направляется на консультацию в медико-генетический центр. Врач-генетик проводит первичную постановку диагноза, чаще формируя дифференциально-диагностический ряд. Окончательный диагноз ставится только по результатам молекулярно-генетического или другого специфического лабораторного исследования.

Однако редкая распространенность данных заболеваний является причиной частых врачебных ошибок [7] и несвоевременной постановки точного диагноза. Так, в Китае время между первым визитом к врачу по поводу симптомов, связанных с МПС, и окончательным диагнозом, составило 9,42 мес., без существенной разницы между типами МПС [8]. Это обусловлено рядом таких причин, как:

- недостаток у многих врачей-генетиков личного опыта по диагностике большинства нозологий на долабораторном этапе [9];
- обширная вариативность проявлений клинических форм (фенотипический полиморфизм);
- нечеткость клинических проявлений, которая определяется возрастом манифестации заболевания и степенью выраженности признаков.

Это является основанием для поиска диагноза в расширенном круге нозологий. Важность точной и ранней постановки диагноза обусловлена тем, что для ряда клинических форм ЛБН существует специфический метод лечения – ферментозаместительная терапия. Это способствует предотвращению прогрессирования патологического процесса и возникновения необратимых изменений.

В то же время методы искусственного интеллекта (ИИ) позволяют обеспечить эффективную поддержку врачебных решений. С их помощью можно значительно улучшить процесс дифференциальной диагностики,

анализируя данные пациентов и предоставляя врачам список потенциальных диагнозов [10]. Это может помочь врачам рассмотреть более широкий спектр потенциальных диагнозов и выявить редкие или сложные состояния, трудные для диагностики.

Таким образом, поддержка врачебного решения в процессе дифференциальной диагностики ЛБН с использованием интеллектуальной системы является актуальной задачей. Средством выбора явилось создание системы на основе экспертных знаний, так как крайне низкая популяционная частота наследственной патологии исключает применение других технологий ИИ.

Материал и методы

Материалом для клинической апробации были ретроспективные данные 92 пациентов с лабораторно и / или молекулярно-генетически подтвержденным диагнозом МПС, МЛП или ГЗ. Критериями включения были: возраст от 0 до 18 лет, направление педиатром на консультацию в специализированное учреждение врачом-генетиком в связи с подозрением на наличие наследственного генетически обусловленного заболевания из группы ЛБН в связи с выявленным сочетанием характерных признаков: грубые черты лица и / или воронкообразная грудная клетка, и / или килевидная грудная клетка, и / или «вишневая косточка» на сетчатке, и / или когтистая лапа, и / или скафоцефалия, и / или макроцефалия. Критерием невключения было подтверждение у пациента иного основного заболевания (ввиду наличия другого генетического дефекта). Выборка содержала 92 деперсонифицированные выписки из электронных медицинских карт (ЭМК) пациентов (МПС – 78, МЛП – 10, ГЗ – 4). Выборка была разделена на 2 части: валидационную и верификационную. Валидация проведена на данных 54 пациентов из Научно-исследовательского клинического института педиатрии и детской хирургии им. акад. Ю.Е. Вельтищева (Москва) – 45 МПС и 9 МЛП, верификация на данных из МГЦ Московского областного научно-исследовательского клинического института им. М.Ф. Владимирского (Москва) – 15 МПС; Клинико-диагностического центра «Охрана здоровья матери и ребенка» (Екатеринбург) – 16 МПС, 1 МЛП и 3 ГЗ; Университетской клиники Института педиатрии Приволжского исследовательского медицинского университета (Нижний Новгород) – 2 МПС и 1 ГЗ. Характеристика набора данных представлена в таблице 1.

Целью валидации было проверить систему на реальных данных пациентов из одной медицинской организации, а цель верификации – оценить стабильность выдаваемых решений на данных из других учреждений трех регионов России.

Выборки, сформированные методом сплошного отбора в соответствии с целью исследования, как видно из таблицы 1, характеризовались крайне низким количеством пациентов с МЛП и ГЗ, что обусловлено их редким поступлением в специализированные стационары (однако подавляющее большинство из них было правильно распознано системой).

Обезличивание обеспечивалось заменой ФИО в выписках из ЭМК на идентификаторы. Слепой метод использовался при клинической апробации (валидации и верификации) системы, выписки из ЭМК, поступающие для компьютерной диагностики, не включали диагнозы пациентов.

Методы инженерии знаний [11] были применены для

Таблица 1. Характеристика выборок

Table 1. Sample characteristics

Параметры	Валидация	Верификация
Количество случаев	54	38
Возраст (мин. – макс., ср.)	0 лет 9 мес. – 17 лет 6 мес. (7 лет 4 мес.)	0 лет 2 мес. – 17 лет 11 мес. (4 года 10 мес.)
Пол (М/Ж)	40/14	25/13
Кол-во МПС	45	33
Кол-во МЛП	9	1
Кол-во ГЗ	0	4

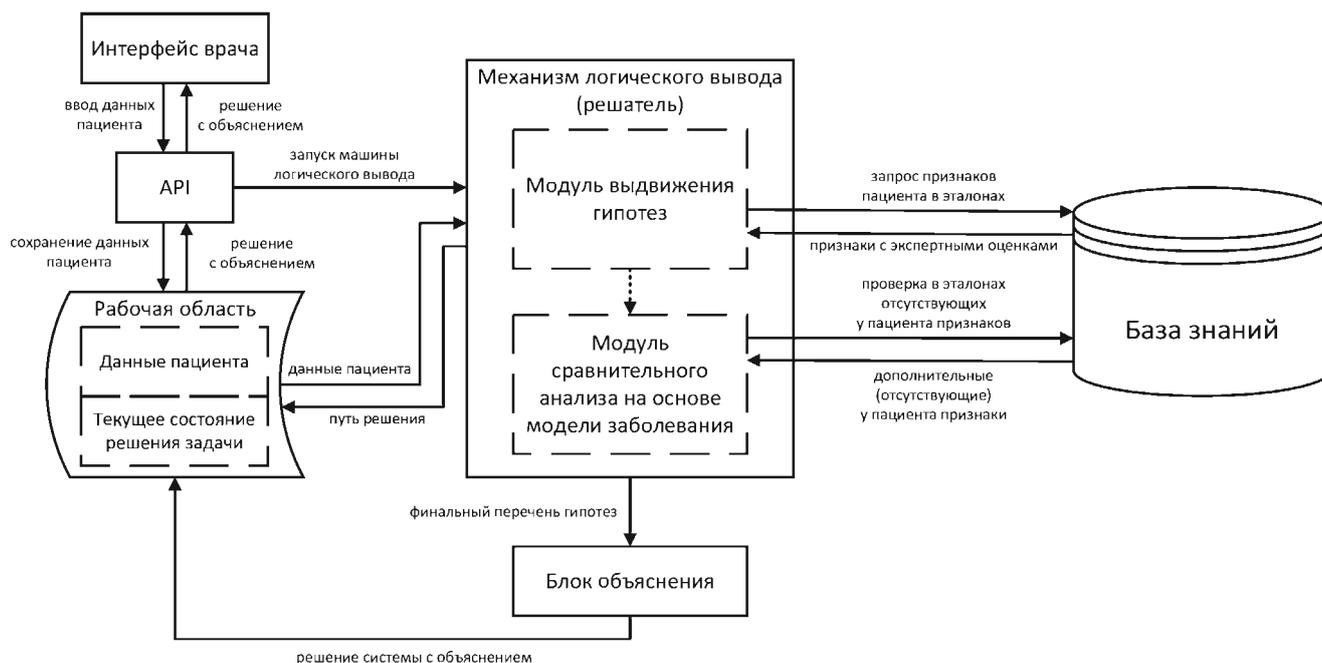


Рисунок. Структурно-функциональная схема системы ГенДиЭС
Figure. Structural and functional diagram of the GenDiES system

извлечения и структуризации знаний. Для оценки значимости каждого признака каждой клинической формы в каждом из 4 возрастных периодов использовались факторы уверенности в модифицированной стэнфордской модели [12] и коэффициент для модальности [13], указывающий на диагностическую важность признака.

Для формализации знаний был применен матричный подход [14], позволивший построить матрицу «болезни – признаки», являющуюся основой базы знаний.

Предыдущая версия системы ГенДиЭС была реализована на базе облачной платформы IACPaaS (<https://iasraas.dvo.ru/>), в связи с чем формализация знаний осуществлялась с использованием онтологического представления. При модификации архитектуры базы знаний осуществлен переход к матричному представлению знаний. Доработана шкала модальности признака. Для принятия решений реализован оригинальный алгоритм, учитывающий условно избыточные признаки [15].

Информационное и программное обеспечение (СУБД PostgreSQL – для хранения базы знаний, язык программирования Java – для программной реализации системы, фреймворк Bootstrap – для отображения интерфейсных форм) позволило реализовать архитектуру интеллектуальной системы.

Результаты

Новая версия экспертной системы ГенДиЭС для дифференциальной диагностики 30 клинических форм ЛБН на долабораторном этапе, в отличие от описанной в [16], реализована как самостоятельное web-приложение. Структурно-функциональная схема системы представлена на рисунке.

Система ГенДиЭС состоит из модулей, которые взаимодействуют между собой. Поступление в систему данных нового пациента инициирует работу программного модуля (API), который выполняет функцию управляющего контроллера, запуская работу механизма логического вывода (решателя). Введенные врачом данные передаются в рабочую область системы на период активной сессии. Из рабочей области данные пациента поступают в запущенный механизм логического вывода, которая выдвигает и сопоставляет поступившую информацию с эталонными описаниями заболеваний в базе знаний. Решатель также осуществляет поиск в базе знаний признаков, отсутствовавших у пациента, но указанных в эталонах, что позволяет информировать врача о полной клинической картине конкретной нозологической формы. Финальный ранжированный перечень гипотез с характеризующими их признаками передается в блок объяснений,

где осуществляется генерация объяснения по заданным шаблонам. Решение системы с объяснением через модуль рабочей области передается на интерфейс врача.

Система ГенДиЭС предполагает работу как через собственный web-интерфейс системы, так и на основе интеграции с медицинскими информационными системами.

Рассмотрим работу отдельных блоков системы в процессе ее использования врачом.

Интерфейс врача-пользователя

Точкой входа в систему является интерфейс врача-пользователя. Он обеспечивает возможность ввода анкетной и клинической информации о пациенте. С целью обеспечения информационной безопасности персональная информация пациента (фамилия, имя и отчество) не запрашиваются, врачу предлагается указать идентификатор пациента в ЭМК или любой иной используемый в медицинской организации кодификатор.

После ввода анкетной информации осуществляется выбор признаков, которые были обнаружены у пациента при осмотре. Он осуществляется из встроенного в систему иерархического справочника, реализованного в форме разноуровневого тезауруса, соответствующего логической структуре протокола осмотра пациента и содержащего 35 признаков. Интерфейсная форма содержит ряд логических проверок введенной врачом информации. При успешном прохождении проверок введенная о пациенте информация передается в рабочую область системы.

Рабочая область

В рабочей области, наряду с исходными данными пациента, хранится промежуточная информация решаемой в текущий момент задачи: выдвинутые и отклоненные гипотезы, сопровождающиеся объяснениями о причинах предлагаемых решений. Промежуточные данные системы необходимы для возможности отслеживания пути принятия решения. Кроме того, реализованный модуль рабочей области позволяет вернуться и скорректировать введенную о пациенте информацию, например, в случае необходимости дополнительно отметить признак, который изначально по каким-то причинам был пропущен. Из рабочей области информация о пациенте передается в блок принятия решений – базу знаний с решателем.

База знаний

Правила для базы знаний формировались на основе структурированной информации из литературных источников, уточнявшейся затем экспертами, которые корректировали описания признаков и отбирали те, что были релевантны для различных ЛБН. Признаки дополнялись 3 экспертными оценками: коэффициентом модальности (важности для диагностики) M_i , факторами уверенности манифестации m_i и выраженности s_i . Каждая клиническая форма описывалась для 4 возрастных периодов: до 1 года, от 1 года до 3 лет включительно, от 4 до 6 лет включительно, 7 лет и старше.

Для экспертных оценок использовались следующие шкалы:

- Модальность (M_i) принимает одно из следующих значений $\{-1;0;1;2;4;5\}$, где:
 - -1 означает, что признак не может наблюдаться при данной клинической форме,
 - 0 – признак не специфичен для данного заболевания,
 - 1 – признак носит очень редкий характер при

данном заболевании,

- 2 – признак является второстепенным,
- 4 – признак является необходимым,
- 5 – признак является главным.

Манифестация (m_i) определяется в интервале от -1 до +1, где:

- -1 означает, что признак не может манифестировать по физиологическим причинам,
- значение в интервале от 0 до 1 с шагом 0,1 – степень уверенности в манифестации.
- Выраженность (s_i) определяется в интервале $[0; 10]$, где:
 - 0 – признак не проявляется,
 - значение в интервале $[1; 10]$ с шагом 1 есть степень выраженности признака.

Структурированные знания были формализованы в виде матрицы «болезни – признаки», где строки – это симптомы, а столбцы – экспертные оценки в 4 возрастных периодах по каждой из клинических форм заболеваний. Структура и пример заполнения матрицы для МПС VII типа (синдром Слая) представлены в таблице 2.

Всего экспертами были выставлены 12 600 оценок, характеризующих проявления (модальность, манифестация, выраженность) 35 признаков для 30 клинических форм МПС, ГЗ и МЛП по четырем возрастным периодам. На этой основе были сформированы эталонные описания каждой клинической формы диагностируемых системной заболеваний.

Решатель

В решателе системы происходит выдвижение и отбор гипотез, т. е. формируется дифференциально-диагностический ряд, который поступает к врачу для окончательной диагностики. Для этого в системе используется алгоритм, в основу которого положен последовательный перебор гипотез в пространстве рассматриваемых диагнозов. На 1-м этапе модуль выдвижения гипотез о диагнозах осуществляет сопоставление признаков пациента с описаниями клинических форм ЛБН в базе знаний с учетом возрастной группы, после чего осуществляется группировка признаков пациента в соответствии с их значимостью для каждой из гипотез. При наличии признаков «за» осуществляется выдвижение гипотез, а при наличии признаков «против» или при отсутствии признаков «за» гипотеза отклоняется. Выдвинутые гипотезы группируются в упорядоченные «кластеры» по количеству признаков вне эталона. На 2-м этапе модуль ранжирования гипотез последовательно осуществляет оценку сходства нового случая с эталонными описаниями нозологических форм ЛБН, ранжирование гипотез внутри «кластеров» по уменьшению сходства с эталонами, формирование единого ранжированного пространства гипотез.

Ранжированный перечень гипотез (дифференциально-диагностический ряд) передается в блок объяснения.

Блок объяснения

В качестве объяснения отобранных гипотез система предоставляет следующие признаки:

- Признаки пациента, подтверждающие гипотезу;
- диагностически значимые проявления заболевания, которые в высокой степени характерны для определенной гипотезы;
- второстепенные проявления заболевания, которые могут встречаться, но не являются специфичными

Таблица 2. Матрица «болезни – признаки» (фрагмент)

Table 2. Matrix “diseases – features” (fragment)

Название признака	Мукополисахаридоз VII типа (синдром Слая)											
	до 1 года			1–3 года			4–6 лет			7 лет и старше		
	<i>Mi</i>	<i>mi</i>	<i>si</i>	<i>Mi</i>	<i>mi</i>	<i>si</i>	<i>Mi</i>	<i>mi</i>	<i>si</i>	<i>Mi</i>	<i>mi</i>	<i>si</i>
Задержка роста	2	0,2	1	2	0,3	4	2	0,5	5	2	0,7	7
Помутнение роговицы	0	–1	0	4	0,3	2	4	0,4	4	4	0,5	7

для гипотезы;

– редкие проявления заболевания, которые описаны в некоторых литературных источниках на уровне случайной находки.

– Признаки пациента, не подтверждающие, но и не отвергающие гипотезу – проявления заболевания, отсутствующие в эталонных описаниях, которые могут указывать на атипичность клинического случая, индивидуальную особенность пациента или могут быть проявлением сопутствующей болезни.

– Признаки, которые не были отмечены у пациента, но их наличие могло бы повысить уверенность в данной гипотезе.

Каждая гипотеза сопровождается справочной информацией, которая включает: накапливаемый субстрат, дефицит фермента, ген, в котором произошла мутация, локализацию гена, код OMIM, код МКБ-10.

Финальный ранжированный перечень гипотез, дополненный объяснениями для каждого заболевания, передается в рабочую область системы. Далее решение системы поступает на интерфейс пользователя. На нем отображается анкетная информация о пациенте и решение системы с объяснением. Врач при необходимости может сохранить его в файл pdf.

Результаты клинической апробации системы

Дифференциально-диагностический ряд, выдаваемый экспертной системой ГенДиЭС, включает до 5 нозологических форм, как это принято в большинстве аналогичных систем [17, 18].

Клиническая апробация проводилась на валидационной и верификационной выборках рассматриваемых 30 лизосомных болезней.

Данные пациентов последовательно загружались в систему и фиксировались выдаваемые решения. При валидации лабораторно подтвержденный диагноз входил в диагностический ряд из 5 гипотез в 47 случаях из 54. При верификации системы входение в топ-5 было зафиксировано в 34 случаях из 38.

На основании этих результатов были рассчитаны оценки точности экспертной системы, которые составили 0,87 (95% ДИ [0,75; 0,95]) и 0,90 (95% ДИ [0,75; 0,97]) для валидации и верификации соответственно.

Результаты клинической апробации показывают воспроизводимость результатов и стабильность работы системы на данных из различных медицинских организаций.

Обсуждение

Проблема своевременной диагностики редких болезней является глобальной. Несмотря на это, в мире насчитывается небольшое количество систем поддержки принятия решений по дифференциальной диагностике на первом (долабораторном) этапе постановки диагноза

на основе фенотипических проявлений болезни. Все эти системы ориентированы на формирование ограниченного круга диагностических гипотез. Сравнение с аналогами можно провести среди современных систем, сведения о точности диагностики для которых есть в открытых источниках: немецкой Ada DX [17], испанской Rare Disease Discovery (RDD) [18] и британской Face2Gene [19, 20].

Экспертная система Ada DX [17] для редких заболеваний показала точность 0,54 по результатам тестирования при условии входения подтвержденного диагноза в перечень из 5 болезней. Система RDD [18] для орфанных болезней показала точность 0,60 по результатам проведенного тестирования при выделении дифференциального ряда из 10 гипотез, при расширении до 50 гипотез точность системы возросла до 0,80. В системе Face2Gene даже при очень редких генетических заболеваниях получена точность 0,75 для 3 болезней дифференциального ряда и 0,83 для 10 гипотез [19, 20].

В новой версии системы ГенДиЭС при валидации и верификации по фенотипическим проявлениям показала достаточно высокую точность, равную 0,87 и 0,90 соответственно при условии входения верного диагноза в ряд из 5 гипотез, что превышает точность рассмотренных современных зарубежных систем. Данный результат получен в процессе верификации в нескольких медицинских организациях, что демонстрирует воспроизводимость на данных, полученных из организаций разных регионов, и сопоставим с аналогичными зарубежными системами, приведенными выше.

Выводы

Разработана новая версия системы ИИ ГенДиЭС для долабораторной диагностики наследственных ЛБН для 30 клинических форм МПС, МЛП и ГЗ в 4 возрастных периодах. Алгоритм принятия решений обеспечивает формирование ранжированного перечня диагнозов – упорядоченного дифференциально-диагностического ряда. Блок объяснения позволяет врачу интерпретировать решение системы на основе признаков пациента, использованных системой для выдвижения соответствующей гипотезы.

Клиническая апробация системы ГенДиЭС на реальных данных больных с подтвержденным диагнозом из 4 медицинских организаций демонстрирует высокую точность долабораторной дифференциальной диагностики для ряда из 5 гипотез: 0,87 – при валидации и 0,90 – при верификации.

Программная реализация ГенДиЭС в виде web-приложения обеспечивает врачам распределенный доступ к системе с любого устройства, подключенного к сети интернет или к серверу медицинской организации, включая возможность интеграции с медицинской информационной системой.

Литература / References

1. Richter T., Nestler-Parr S., Babela R., Khan Z.M., Tesoro T., Molsen E. et al. Rare diseases terminology and definitions – A systematic global review: Report of the ISPOR Rare Disease Special Interest Group. *Value Health*. 2015;18(6):906–914. <https://doi.org/10.1016/j.jval.2015.05.008>
2. Балкунова Ю.Н., Кочергина Е.А., Базанова Н.А. Клинический пример: случай муколипидоза II/IIIА типа у ребенка. *Пермский медицинский журнал*. 2023;40(4):135–140. <https://doi.org/10.17816/pmj404135-140>
3. Балкунова Ю.Н., Кочергина Е.А., Базанова Н.А. Clinical case: type II/IIIА mucopolidosis in a child. *Perm Medical Journal*. 2023;40(4):135–140. (In Russ.). <https://doi.org/10.17816/pmj404135-140>
3. Горбунова В.Н. Наследственные болезни обмена. Лизосомные болезни накопления. *Педиатр*. 2021;12(2):73–83. <https://doi.org/10.17816/PED12273-83>
3. Gorbunova V.N. Congenital metabolic diseases. Lysosomal storage diseases. *Pediatrician (St. Petersburg)*. 2021;12(2):73–83. (In Russ.). <https://doi.org/10.17816/PED12273-83>
4. Borges P., Pasqualim G., Giugliani R., Matte U. Estimated prevalence of mucopolysaccharidoses from population-based exomes and genomes. *Orphanet J. Rare Dis*. 2020;15:324. <https://doi.org/10.1186/s13023-020-01608-0>
5. Захарова Е.Ю., Байдакова Г.В., Михайлова С.В., Пчелина С.Н., Краснополская К.Д. Лизосомные болезни накопления: руководство для врачей. М: ГЭОТАР-Медиа; 2021:424. <https://doi.org/10.33029/9704-6321-5-LAD-2021-1-424>. ISBN 978-5-9704-6321-5.
5. Zakharova E.Yu., Baidakova G.V., Mikhailova S.V., Pchelina S.N., Krasnopol'skaya K.D. Lysosomal Storage Diseases: A Guide for Clinicians. Moscow: GEOTAR-Media; 2021:424. <https://doi.org/10.33029/9704-6321-5-LAD-2021-1-424>. ISBN 978-5-9704-6321-5.
6. Назаренко Л.П., Назаренко М.С. Особенности раннего проявления лизосомных болезней накопления. *Медицинская генетика*. 2013;12(9):20–23. EDN: TJBVIV
6. Nazarenko L.P., Nazarenko M.S. THE early symptoms of lysosomal storage diseases. *Medical Genetics*. 2013;12(9):20–23. (In Russ.). EDN: TJBVIV
7. Кравчук Ж.П., Румянцева О.А. Орфанные заболевания: определение, проблемы, перспективы. *Проблемы здоровья и экологии*. 2013;(4):7–11. <https://doi.org/10.51523/2708-6011.2013-10-4-1>
7. Kravchuk Z.P., Romyantseva O.A. Orphan diseases: diagnosis, problems, prospects. *Health and Ecology Issues*. 2013;(4):7–11. (In Russ.). <https://doi.org/10.51523/2708-6011.2013-10-4-1>
8. Kang Q., Fang Y., Yang Y., Li D., Zheng L., Chen X. et al. Health service utilization, economic burden and quality of life of patients with mucopolysaccharidosis in China. *Orphanet J. Rare Dis*. 2024;19:324. <https://doi.org/10.1186/s13023-024-03333-4>
9. Шашель В.А., Фирсова В.Н., Трубилина М.М., Подпорина Л.А., Фирсов Н.А. Орфанные заболевания и связанные с ними проблемы. *Медицинский вестник Юга России*. 2021;12(2):28–35. <https://doi.org/10.21886/2219-8075-2021-12-2-28-35>
9. Shashel V.A., Firsova V.N., Trubilina M.M., Podporina L.A., Firsov N.A. Orphan diseases and associated problems. *Medical Herald of the South of Russia*. 2021;12(2):28–35. (In Russ.). <https://doi.org/10.21886/2219-8075-2021-12-2-28-35>
10. Krishnan G., Singh S., Pathania M., Gosavi S., Abhishek S., Parchani A., Dhar M. Artificial intelligence in clinical medicine: catalyzing a sustainable global healthcare paradigm. *Frontiers in Artificial Intelligence*. 2023;6:1227091. <https://doi.org/10.3389/frai.2023.1227091>
11. Гаврилова Т.А., Кудрявцев Д.В., Муромцев Д.И. Инженерия знаний. Модели и методы: Учебник. СПб.: Лань; 2023:324. ISBN: 978-5-507-46580-4.
12. Кобринский Б.А. Триединство факторов уверенности в задачах медицинской диагностики. *Искусственный интеллект и принятие решений*. 2018;(2):62–72. <https://doi.org/10.14357/20718594180205>
12. Kobrinskii B.A. Certainty factors triunity in the medical diagnostics tasks. *Artificial Intelligence and Decision Making*. 2018;(2):62–72. (In Russ.). <https://doi.org/10.14357/20718594180205>
13. Грибова В.В., Петряева М.В., Шалфеева Е.А. Облачный сервис поддержки принятия диагностических решений в гастроэнтерологии. *Врач и информационные технологии*. 2019;(3):65–71.
13. Gribova V.V., Petryaeva M.V., Shalfeeva E.A. Cloud decision support service for diagnosis in gastroenterology. *Medical doctor and information technologies*. 2019;(3):65–71. (In Russ.).
14. Орехова Т.Ф., Кружилина Т.В., Неретина Т.Г. Матричный подход к описанию педагогических процессов в научных педагогических исследованиях. *Бизнес. Образование. Право*. 2020;(2):301–308. <https://doi.org/10.25683/VOLBI.2020.51.213>
14. Orekhova T.F., Kruzhillina T.V., Neretina T.G. Matrix approach to the description of pedagogical processes in scientific pedagogical research. *Business. Education. Law*. 2020;(2):301–308. (In Russ.). <https://doi.org/10.25683/VOLBI.2020.51.213>
15. Благодосклонов Н.А., Кобринский Б.А. Принятие диагностических решений экспертной системой в условиях неполноты или избыточности данных. Интегрированные модели и мягкие вычисления в искусственном интеллекте. Сборник научных трудов XII Международной научно-практической конференции (ИММБ-2024, Коломна, 14–17 мая 2024 г.). В 2-х т. Т. 1. Смоленск: Универсум; 2024:55–63. EDN: WKNXQX
15. Blagoslkonov N.A., Kobrinskii B.A. Decision making in conditions of incomplete or redundancy data. Integrated models and soft computing in artificial intelligence. Proceedings of the XII International Scientific and Practical Conference (IMSC-2024, Kolomna, May 14–17, 2024). In 2 vol. Vol. 1. Smolensk: Universum; 2024:55–63. (In Russ.). EDN: WKNXQX
16. Благодосклонов Н.А., Кобринский Б.А. Дифференциальная диагностика наследственных болезней обмена с использованием системы на основе экспертных знаний. *Сибирский журнал клинической и экспериментальной медицины*. 2020;35(4):71–78. <https://doi.org/10.29001/2073-8552-2020-35-4-71-78>
16. Blagoslkonov N.A., Kobrinskii B.A. Differential diagnosis of hereditary metabolic diseases using the expert knowledge-based system. *Siberian Journal of Clinical and Experimental Medicine*. 2020;35(4):71–78. (In Russ.). <https://doi.org/10.29001/2073-8552-2020-35-4-71-78>
17. Ronicke S., Hirsch M.C., Türk E., Larionov K., Tientcheu D., Wagner A.D. Can a decision support system accelerate rare disease diagnosis? Evaluating the potential impact of Ada DX in a retrospective study. *Orphanet J. Rare Dis*. 2019;(14):69. <https://doi.org/10.1186/s13023-019-1040-6>
18. Alves R., Piñol M., Vilaplana J., Teixidó I., Corella J., Comas J. et al. Computer-assisted initial diagnosis of rare diseases. *PeerJ*. 2016;4:e2211. <https://doi.org/10.7717/peerj.2211>
19. Pantel J.T., Zhao M., Mensah M.A., Hajji N., Hsieh T.-C., Hanani Y. et al. Advances in computer-assisted syndrome recognition by the example of inborn errors of metabolism. *J. Inherit. Metab. Dis*. 2018;41(3):533–539. <https://doi.org/10.1007/s10545-018-0174-3>
20. Carrer A., Romaniello M.G., Calderara M.L., Mariani M., Biondi A., Selicorni A. Application of the Face2Gene tool in an Italian dysmorphological pediatric clinic: Retrospective validation and future perspectives. *Am. J. Med. Genet. Part A*. 2024;194(3):e63459. <https://doi.org/10.1002/ajmg.a.63459>

Информация о вкладе авторов

Кобринский Б.А. – разработка концепции и дизайна исследования, анализ и интерпретация результатов, окончательное утверждение рукописи для публикации, обоснование рукописи и проверка критически важного интеллектуального содержания.

Благосклонов Н.А. – разработка дизайна исследования, сбор данных, анализ данных и интерпретация результатов, подготовка текста рукописи.

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

Information on author contributions

Kobrinskii B.A. – study concept and design, analysis and interpretation of results, final approval for publication, justification of the manuscript and verification of critical intellectual content.

Blagoslkonov N.A. – study design, data collection and analysis, interpretation of the results, text preparation.

Conflict of interest: the authors declare no conflict of interest.

Сведения об авторах

Кобринский Борис Аркадьевич, д-р мед. наук, профессор, заведующий отделом систем интеллектуальной поддержки принятия решений, ФИЦ ИУ РАН; профессор кафедры медицинской кибернетики и информатики им. С.А. Гаспаряна, Институт биомедицины (МБФ), РНИМУ им. Н.И. Пирогова Минздрава России (Пироговский Университет), Москва, Россия, <http://orcid.org/0000-0002-3459-8851>.

E-mail: kba_05@mail.ru.

Благосклонов Николай Алексеевич, научный сотрудник, отдел систем интеллектуальной поддержки принятия решений, ФИЦ ИУ РАН; старший преподаватель, кафедра медицинской кибернетики и информатики им. С.А. Гаспаряна, Институт биомедицины (МБФ), РНИМУ им. Н.И. Пирогова Минздрава России (Пироговский Университет), Москва, Россия, <http://orcid.org/0000-0002-5293-8469>.

E-mail: nblagosklonov@frccsc.ru.

 **Кобринский Борис Аркадьевич**, e-mail: kba_05@mail.ru.

Поступила 02.06.2025;
рецензия получена 05.06.2025;
принята к публикации 09.06.2025.

Information about the authors

Boris A. Kobrinskii, Dr. Med. (Sci.), Professor, Head of the Department of Intelligent Decision Support Systems, Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences; Professor, S.A. Gasparyan Department of Medical Cybernetics and Informatics of the Institute of Biomedicine (MBF), Pirogov Russian National Research Medical University, Moscow, Russia, <http://orcid.org/0000-0002-3459-8851>.

E-mail: kba_05@mail.ru.

Nikolay A. Blagosklonov, Research Scientist, Department of Intelligent Decision Support Systems, Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences; Senior Lecturer, S.A. Gasparyan Department of Medical Cybernetics and Informatics of the Institute of Biomedicine (MBF), Pirogov Russian National Research Medical University, Moscow, Russia, <http://orcid.org/0000-0002-5293-8469>.

E-mail: nblagosklonov@frccsc.ru.

 **Boris A. Kobrinskii**, e-mail: kba_05@mail.ru.

Received 02.06.2025;
review received 05.06.2025;
accepted for publication 09.06.2025.