



<https://doi.org/10.29001/2073-8552-2020-35-4-71-78>

УДК 616-008.9-056.7-079.4:004.891

Дифференциальная диагностика наследственных болезней обмена с использованием системы на основе экспертных знаний

Н.А. Благосклонов, Б.А. Кобринский

Федеральный исследовательский центр «Информатика и управление» Российской академии наук, 117312, Российская Федерация, Москва, пр. 60-летия Октября, 9

Аннотация

Цель исследования: создание компьютерной системы поддержки принятия решений с использованием экспертных знаний для диагностики редких наследственных болезней, целесообразность которой определяется трудностью их идентификации на долабораторном этапе.

Материал и методы. В качестве материала исследования использованы описания клинической картины лизосомных болезней накопления (ЛБН) из литературных источников. Примененные методы включали извлечение знаний, экспертные оценки, квантование возрастных интервалов, прикладные интеллектуальные сервисы для формирования базы знаний.

Результаты. Результаты исследования включают построение моделей комплексной оценки признака и интегральной оценки заболевания, на основе которых реализуется алгоритм сопоставительного анализа для оценки каждой из гипотез, выдвинутых системой. Результаты тестирования прототипа созданной экспертной системы на контрольной выборке больных мукополисахаридозами (МПС) показали эффективность 90%.

Обсуждение. В обсуждении рассмотрены несколько диагностических систем, показано их отличие от системы, представленной в настоящей работе.

Заключение. Обобщены итоги разработки интеллектуальной системы на основе знаний для диагностики ЛБН и указано на перспективы ее развития.

Ключевые слова:	компьютерная диагностика наследственных болезней обмена, экспертная система, коэффициенты модальности признаков, факторы уверенности, лизосомные болезни накопления, долабораторная диагностика.
Конфликт интересов:	авторы заявляют об отсутствии конфликта интересов.
Прозрачность финансовой деятельности:	никто из авторов не имеет финансовой заинтересованности в представленных материалах или методах. Работа выполнена в рамках государственного задания.
Для цитирования:	Благосклонов Н.А., Кобринский Б.А. Дифференциальная диагностика наследственных болезней обмена с использованием системы на основе экспертных знаний. <i>Сибирский журнал клинической и экспериментальной медицины</i> . 2020;35(4):71–78. https://doi.org/10.29001/2073-8552-2020-35-4-71-78 .

Differential diagnosis of hereditary metabolic diseases using the expert knowledge-based system

Nikolay A. Blagosklonov, Boris A. Kobrinskii

Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 9, 60-th October Anniversary pr., Moscow, 117312, Russian Federation

Abstract

Aim. The aim of the study was to create a computer decision support system using expert knowledge for the diagnosis of rare hereditary diseases due to the difficulty of their identification at the pre-laboratory stage.

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

Material and Methods. Descriptions of the clinical picture of lysosomal storage diseases from literature sources were used as the research material. The methods included knowledge extraction, expert assessments, quantization of age intervals, and applied intelligent services to form a knowledge base.

Results. The results of the study include the construction of models for a complex assessment of a sign and an integral assessment of a disease, on the basis of which the comparative analysis algorithm is implemented to assess each of the hypotheses put forward by the system. The results of testing the prototype of the created expert system on a control sample of patients with mucopolysaccharidosis showed the efficiency of 90%.

Discussion. In the discussion, several diagnostic systems are considered and their distinction from the system, presented in this work, is shown.

Conclusion. The results of the development of intelligent system based on knowledge for the diagnosis of lysosomal storage diseases are summarized and the perspectives for its development are highlighted.

Keywords:	computer diagnostics of hereditary metabolic diseases, expert system, modality coefficients of signs, certainty factors, lysosomal storage diseases, pre-laboratory diagnostics.
Conflict of interest:	the authors do not declare a conflict of interest.
Financial disclosure:	no author has a financial or property interest in any material or method mentioned. The work was performed as part of a state assignment.
For citation:	Blagosklonov N.A., Kobrinskii B.A. Differential diagnosis of hereditary metabolic diseases using the expert knowledge-based system. <i>The Siberian Journal of Clinical and Experimental Medicine</i> . 2020;35(4):71–78. https://doi.org/10.29001/2073-8552-2020-35-4-71-78 .

Введение

Принципиальные звенья патогенеза генных болезней можно представить в виде следующей последовательности: мутантный аллель, патологический первичный продукт, цепь биохимических процессов, поражение клеток, органов, организма [1]. Унаследованные дефекты ферментов являются причиной накопления недеградированных метаболитов. В зависимости от накапливаемого метаболита среди лизосомных болезней накопления (ЛБН) выделяют мукополисахаридозы, сфинголипидозы, муколипидозы и другие [2]. Так, при мукополисахаридозах (МПС) дефицит специфического лизосомного фермента вызывает прогрессирующее накопление гликозаминогликанов, что приводит к дисфункции клеток, тканей и многих органов [3].

Несмотря на относительную редкость отдельных нозологических форм ЛБН, при дифференциальной диагностике патологических состояний необходимо учитывать их проявление уже в неонатальном периоде, хотя встречаются и формы для взрослых [4, 5]. Хорошее знание признаков, характеризующих клиническую картину этих заболеваний, может помочь уменьшить ошибки диагностики, содействовать раннему выявлению патологии и назначению специальной терапии. Ранняя диагностика особенно важна, если доступно лечение, модифицирующее проявления заболевания [6, 7], что можно отметить в отношении ряда лизосомных болезней. Однако большая гетерогенность клинических проявлений как между, так и внутри каждого типа и подтипа этих заболеваний вызывает значительные трудности для их идентификации. Не вызывает сомнения, что существует острая необходимость в улучшении диагностики наследственных болезней [8]. В целом их идентификация представляет собой когнитивную проблему из-за комбинации факторов, которые включают редкость заболеваний и ограничения памяти человеческого мозга [9].

В связи с этим в течение многих лет ведется поиск методов компьютерной диагностики, которые могли бы максимально эффективно оказать врачам поддержку в принятии диагностических решений при наследственной

патологии. Современный анализ показал, что первичная диагностика редких болезней с использованием экспертных систем на основе анализа фенотипических признаков превышает 80% [10]. Ранее полинозологические системы были созданы в Великобритании, Австралии, Франции, России. Но следует отметить, что в настоящее время реально функционирующими и наиболее распространенными остались всего две системы. London Neurogenetics Database and Dysmorphology (LNDD) направлена на помощь в выявлении и классификации у детей дисморфических проявлений различных заболеваний [11]. В системе имеется привязка синдромов к дефектным генам, также для всех синдромов указываются коды OMIM. В настоящее время LNDD распространяется в web-версии (<https://www.face2gene.com/lmd-history/>) через систему Face2Gene [12]. В Murdoch Institute (Австралия) была создана система POSSUM (Pictures of Standard Syndromes and Undiagnosed Malformations), которая обеспечивает дифференциальную диагностику метаболических болезней, множественных пороков развития, тератогенных и хромосомных синдромов [13].

Однако наибольшие проблемы по-прежнему вызывает ранняя долабораторная диагностика наследственных болезней обмена, для которых отсутствуют методы скрининга новорожденных. К ним относятся, в частности, лизосомные болезни. Поэтому поиск в направлении диагностики этой патологии с использованием систем искусственного интеллекта на основе специальных знаний продолжается.

Материал и методы

В качестве материала исследования для прототипа диагностической системы были использованы описания клинической картины ЛБН (МПС, ганглиозидозы, муколипидозы), в первую очередь клинические случаи, представленные в отечественных и зарубежных статьях и монографиях (в общей сложности в 97 литературных источниках), а также в информационных базах Orphanet

(<https://www.orpha.net/consor/cgi-bin/index.php>), Online Mendelian Inheritance in Man – OMIM (<https://omim.org/>), Genetic and Rare Diseases Information Center – GARD (<https://rarediseases.info.nih.gov/>) и Human Phenotype Ontology – HPO (<https://hpo.jax.org/app/>).

Для МПС, рассматриваемых далее в качестве примера в настоящей статье, были отобраны 80 литературных источников. Они включают 6 типов, представленных в виде 15 клинических форм в OMIM.

Материал был разделен на основную (обучающую) и контрольную выборки из литературных источников. Обучающая выборка использовалась для извлечения знаний о ЛБН с последующим их представлением в структурированном виде. Контрольная выборка не применялась для извлечения знаний о заболеваниях, а только для тестирования разработанного прототипа системы на основе 20 публикаций в научных журналах, описывающих клинические случаи из практики (case reports).

В исследовании использовались методы онтологического инжиниринга (извлечение знаний из литературных источников и у экспертов) [14, 15]. Литературные данные анализировались путем семантического, текстологического и лингвистического методов анализа текстов. Экспертные оценки в виде факторов уверенности использовались в модификации, позволяющей оценить более одной характеристики признаков [16].

Платформа облачных вычислений IACPaaS [17] для прикладных интеллектуальных сервисов применялась для формирования проблемно-ориентированной базы знаний, обеспечивающей принятие решений.

Метод квантования по возрастным интервалам был применен для учета прогрессивности (динамики патологического процесса) ЛБН. В процессе исследования предложенная ранее шкала [18] была частично модифицирована, в результате чего 4 возрастные градации включали: 1-й год жизни, 1–3 года включительно, 4–6 лет включительно, 7 лет и старше. Эти возрастные интервалы учитывались при формировании экспертных оценок.

Специальная шкала была разработана в рамках настоящего исследования для количественной оценки коэффициентов модальности признаков.

Результаты

Разработка системы включала ряд этапов. Вначале описание клинических форм МПС из литературных источников и информационных баз были агрегированы и представлены в структурированной форме в текстологических картах [19]. Для этого различные описания клинических признаков, включая синонимы, приводились к однообразной форме. Когнитолог заносил в карты эту структурированную информацию о наличии признаков, их выраженности и частоте по каждой клинической форме заболевания с указанием возрастной группы, также указывались ссылки на источник. Названия признаков записывались на двух языках (русском, английском).

Для формализованного представления знаний в текстологической карте были разработаны и использовались следующие шкалы:

- шкала присутствия признаков в форме «+» и «–», указывающая на наличие или отсутствие симптома;
- ранговая шкала выраженности признаков, включающая 5 лингвистических оценок: очень сильно выражен, сильно выражен, умеренно выражен, слабо выражен, очень слабо выражен;

– интервальная шкала частоты встречаемости признаков: очень часто (более 80%), часто (в 60–80%), сравнительно часто (от более 30 до 60%), редко (от более 15 до 30%), очень редко (менее 15%).

На следующем этапе разработки системы текстологические карты использовались экспертами. Привлечение специалистов в области наследственных заболеваний было необходимо по ряду причин, которые необходимо учитывать: а) вариации спектра патологических проявлений, б) нечеткость манифестации признаков и динамики нарастания симптоматики, в) фенотипические особенности российской популяции, что обусловлено наличием в зарубежных литературных источниках описаний больных различных этнических групп.

Экспертами был определен необходимый перечень признаков для каждой из клинических форм ЛБН. Для 15 клинических форм МПС, рассматриваемых далее, количество таких признаков составило 22.

Для каждого признака два эксперта совместно определяли три оценки: коэффициент модальности и два фактора уверенности или меры доверия, позволяющие количественно оценить наличие манифестации и определенный уровень выраженности признаков в конкретный возрастной период.

Коэффициент модальности характеризовал диагностическую значимость признаков, которая была представлена тремя градациями, упорядоченными по степени важности: главные, необходимые и второстепенные. Фактор уверенности манифестации признака оценивался в интервале $[-1; 1]$, где «–1» соответствовало невозможности проявления признака у ребенка в определенном возрасте, «0» характеризовал норму (то есть отсутствие патологического проявления), а значения в диапазоне от 0,1 до 1 представляли собой факторы уверенности экспертов в проявлении признака. Для факторов уверенности существовало условие, что сумма отличных от «–1» значений по четырем возрастным группам для каждого признака не может превышать «1». Понятие манифестации к определенному возрасту включало сумму значений текущего и предшествующих факторов уверенности. Соответственно, в каждой из 4 возрастных групп эксперты проставляли фактор уверенности в отношении манифестации в текущем возрасте. Необходимо отметить фундаментальное различие между значениями «–1» и «0». Значение «–1» использовалось экспертами в случаях, когда признак не мог проявиться по физиологическим причинам. Например, широко расставленные зубы у детей в возрасте до 1 года невозможно выявить в связи с тем, что до 1 года прорезаются в среднем только 8 зубов, и судить по ним о межзубных промежутках не вполне корректно. «0» использовался экспертами более гибко, и данному значению соответствовали несколько ситуаций:

- манифестация признака обычно не происходит в силу физиологических возрастных особенностей, однако эксперты допускают, что такое может иметь место в виде исключения;
- манифестация признака в данной возрастной группе не происходит;
- манифестация признака уже завершилась в предыдущем возрастном периоде.

Таким образом, фактор уверенности для манифестации признака количественно характеризовал вероятность патологических изменений, вызванных в том или ином возрастном периоде накоплением гликозаминогликанов в

организме. Интервал [0; 10] аналогичным образом использовался для оценки выраженности проявления признаков.

Экспертные оценки (всего 8 796 для МПС, ганглиозидозов и мукополидозов) представлены в виде матрицы «болезни – признаки», где строки – это симптомы, а столбцы – экспертные оценки в 4 возрастных периодах по каждой из клинических форм заболеваний. В прототипе экспертной системы, реализованном на основе 15 клинических форм МПС, использовано 3 960 экспертных оценок.

При создании прототипа экспертной системы ГенДиЭС (GenDiES) для поддержки принятия диагностических решений при наследственных болезнях обмена формализованные знания по группе МПС были перенесены в облачную платформу IACPaas [17]. Подход к представлению знаний в виде онтологий, используемый в данной платформе, представлен на рисунке 1.

Это фрагмент описания МПС I типа (синдром Гурлер).

На вход диагностической системы подавалась информация о возрастах пациентов и имеющихся признаках заболеваний. Для последующего тестирования полученных гипотез были разработаны модели комплексной оценки признака и интегральной оценки заболевания.

Комплексная оценка признака определяется как произведение коэффициента модальности и факторов уверенности манифестации и выраженности признаков. В комплексной оценке признака использовался фактор уверенности манифестации к определенному возрасту. Данные оценки были рассчитаны для всех признаков в каждом из возрастных периодов каждой клинической формы МПС.

Интегральная оценка заболевания находится как сумма комплексных оценок признаков конкретной клинической формы в каждом из 4 возрастных периодов.

Интегральные оценки, полученные на основе выборки, использованной для «обучения» системы, получили название эталонных. Эталонная оценка рассчитывалась один раз по всем 22 признакам, используемым в диагностической системе для МПС. Персональная оценка, получаемая при дифференциальной диагностике новых случаев на контрольной (тестовой) выборке, рассчитывалась как сумма тех признаков, которые выявлены у пациента. Однако алгоритм сопоставительного анализа интегральных оценок с персональными для каждой из гипотез, выдвинутых системой, учитывает все наличествующие в конкретном случае признаки из возможных 22 для МПС (вне зависимости от отсутствия признаков какой-либо модальности). Кроме того, осуществлялось сопоставление количества признаков, подтверждающих (признаки «за») или опровергающих (признаки «против») каждую из диагностических гипотез. Признак считается опровергающим диагностическую гипотезу, если он не характерен для данной клинической формы.

По результатам проведения ряда экспериментальных тестирований был выбран следующий алгоритм для работы с первично полученными гипотезами.

На первом шаге анализируется количество признаков «за» и «против» в отношении каждой из диагностических гипотез.

На втором шаге подсчитывается, сколько гипотез не имеют признаков «против». На этом процесс мог бы быть закончен при наличии пяти и более гипотез. Однако в связи с нечеткостью времени манифестации признаков при рассматриваемой патологии процесс отбора гипотез продолжается за счет включения имеющих вначале хотя бы один признак

«против» и далее путем увеличения числа контрпризнаков, пока число гипотез не достигнет как минимум пяти.

На третьем шаге осуществляется вычисление персональной интегральной оценки для каждой из отобранных гипотез по присутствующим у пациента признакам. Затем осуществляется сопоставление персональной и эталонной оценок, то есть рассчитывается процент совпадений для отобранных гипотез. Таким образом находится соответствие клинических проявлений у пациента с диагностическими правилами в базе знаний. Если у пациента отсутствуют признаки, важные для определенной клинической формы заболевания, то в таком случае персональная оценка может значительно отличаться от эталонной, чем в случае отсутствия клинических проявлений второстепенных признаков.

На четвертом шаге гипотезы ранжируются по проценту совпадения персональных оценок с интегральной оценкой – от максимального к минимальному, и отбрасывается 5 первых, являющихся наиболее вероятными. В связи с этим следует отметить, что обычно в системах долабораторной дифференциальной диагностики редких заболеваний выдается ряд гипотез, включающий от 3 до 10 диагнозов или полный список по желанию пользователя [9, 10, 20].

Данный алгоритм был протестирован на 20 случаях описаний шести типов МПС из отечественных и зарубежных публикаций.

Распределение форм МПС в тестируемой выборке отображено в таблице 1.

Таблица 1. Распределение форм мукополисахаридозов в тестовой выборке

Table 1. Distribution of mucopolysaccharidosis forms in the test sample

Клиническая форма мукополисахаридоза Clinical form of mucopolysaccharidosis	Количество Amount
Мукополисахаридоз I типа (синдром Гурлер) Mucopolysaccharidosis type I (Hurler syndrome)	3
Мукополисахаридоз I типа (синдром Гурлер – Шейе) Mucopolysaccharidosis type I (Hurler – Scheie syndrome)	1
Мукополисахаридоз II типа (болезнь Хантера) Mucopolysaccharidosis type II (Hunter syndrome)	2
Мукополисахаридоз III типа А (синдром Санфилиппо А) Mucopolysaccharidosis type III A (Sanfilippo syndrome A)	1
Мукополисахаридоз III типа В (синдром Санфилиппо В) Mucopolysaccharidosis type III B (Sanfilippo syndrome B)	1
Мукополисахаридоз III типа С (синдром Санфилиппо С) Mucopolysaccharidosis type III C (Sanfilippo syndrome C)	1
Мукополисахаридоз IV типа А (синдром Моркио А) Mucopolysaccharidosis type IV A (Morquio syndrome A)	2
Мукополисахаридоз IV типа В (синдром Моркио В) Mucopolysaccharidosis type IV B (Morquio syndrome B)	1
Мукополисахаридоз VI типа (Марото – Лами синдром) Mucopolysaccharidosis type VI (Maroteaux – Lami syndrome)	4
Мукополисахаридоз VII типа (синдром Слая) Mucopolysaccharidosis type VII (Sly syndrome)	4

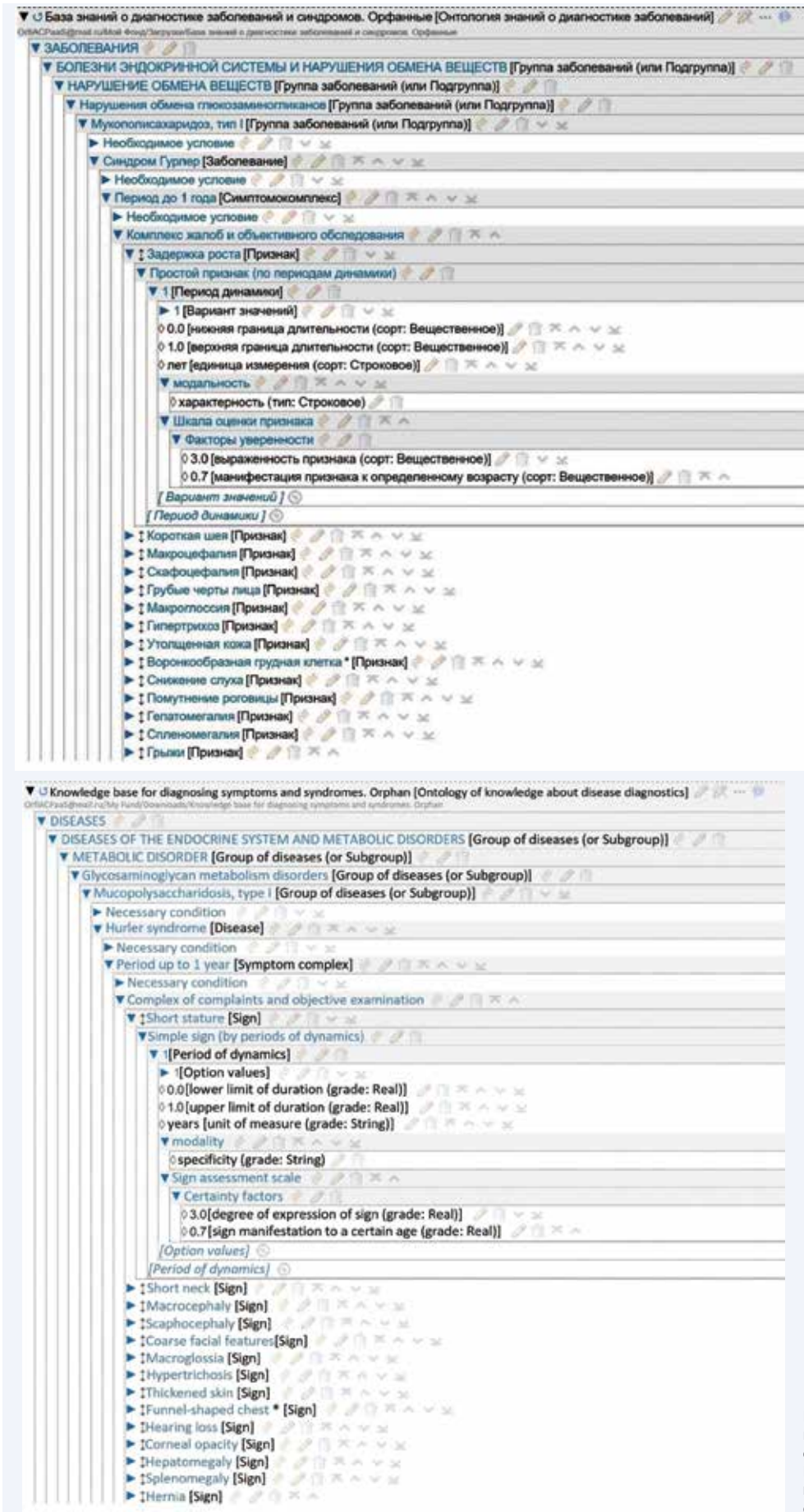


Рис. 1. Пример онтологии мукополисахаридозов на базе облачной платформы IACPaaS
Fig. 1. Example of the mucopolysaccharidoses ontology based on the IACPaaS cloud platform

Необходимо отметить, что подобное распределение МПС не является характерным, так как I и II типы встречаются чаще [21]. Такая выборка обусловлена научным интересом исследователей к менее известному VII типу (синдром Слая), который был обнаружен позднее других.

Эксперименты, осуществленные в процессе тестирования полученных прототипом системы диагностических гипотез, показали наличие сходных результатов для различных типов МПС.

По итогам проведенного эксперимента правильный диагноз среди отобранных первых пяти гипотез присутствовал в 18 случаях, что соответствует точности 90%. Последовательность диагнозов представлена в таблице 2.

Таблица 2. Позиции диагнозов по результатам тестирования прототипа экспертной системы

Table 2. Position of diagnoses on the test results of the expert system prototype

Позиция диагноза в дифференциальном ряду Position in differential row diagnosis	Количество случаев Amount cases
Первый диагноз First diagnosis	6
Второй диагноз Second diagnosis	5
Третий диагноз Third diagnosis	2
Четвертый диагноз Fourth diagnosis	3
Пятый диагноз Fifth diagnosis	2

В то же время отдельное внимание необходимо уделить двум случаям, в которых подтвержденный литературный диагноз не вошел в первые 5 по результатам диагностики с помощью системы, основанной на знаниях. В первом случае рассматривался пациент в возрасте 1 год 1 мес. с диагнозом «мукополисахаридоз VI типа» (синдром Марото – Лами). Однако системой была выдвинута гипотеза, что данная клиническая картина больше соответствует МПС I и II типов. Последующий анализ показал, что это обусловлено необычно ранней манифестацией заболевания при очень невыраженной клинической картине. Во втором случае у пациента в возрасте 2 лет с установленным диагнозом «мукополисахаридоз II типа» (синдром Хантера) верный диагноз оказался на 6-м месте среди выдвинутых гипотез и поэтому не был включен в дифференциально-диагностический ряд. Это объясняется нетипичной клинической картиной случая синдрома Хантера, который характеризовался отсутствием у пациента двух признаков (низкого роста и короткой шеи), существенных для тяжелой формы течения, которая подтверждалась задержкой развития. Нетипичное течение продолжалось, как было отмечено в описании динамики заболевания у этого пациента, когда в возрасте 7 лет рост по-прежнему соответствовал нормальному.

Обсуждение

Прогрессирующий характер ЛБН выдвигает дополнительные требования к системам компьютерной диагностики. Во-первых, это касается учета повозрастной динамики клинической картины заболеваний, во-вторых, – оценки выраженности этих проявлений.

В системе, созданной на основе данных ORPHANET, которые хорошо организованы и тщательно курируются медицинскими экспертами, у 60% пациентов правильно идентифицированное по клиническим признакам заболевание было в первой десятке списка диагнозов при недостаточно полном описании. Этот процент возрастал до 80% для скорректированных описаний клинической картины [10]. Однако следует принять во внимание, что симптомы, отсутствующие в базе данных или напрямую не связанные с предполагаемыми заболеваниями, не подлежали рассмотрению. Также в данной системе отсутствовал учет возраста диагностируемых пациентов, так как в ORPHANET представлены обобщенные (вневозрастные) описания по нозологиям. Ранжирование кандидатов в болезни с помощью статистической модели осуществлялось в семантической сети [20], где применялись метрики подобия для измерения фенотипического сходства между запросами и наследственными заболеваниями, аннотированными с использованием Human Phenotype Ontology. Как отмечают авторы, учет семантических взаимосвязей между терминами относительно эффективен для запросов, содержащих фенотипический шум или неточные клинические описания. Но в этом подходе также не была представлена возрастная динамика заболеваний. Кроме того, в протоколах объяснений выдаваемых диагностических гипотез, в отличие от представленных в настоящей работе, не были указаны подтверждающие признаки и признаки, являющиеся контраргументами.

В немецкой экспертной системе Ada DX [9] при тестировании прототипа верный диагноз среди пяти наиболее подходящих вариантов заболевания был в 53,8% случаев, а в качестве наиболее подходящего варианта заболевания – в 37,6% случаев, что ниже показателей, полученных при тестировании прототипа российской экспертной системы ГенДиЭС, представленной в настоящей публикации. Хотя несомненно интересно, что в ряде случаев Ada DX предоставила правильные рекомендации на ранних стадиях заболевания.

Эффективность созданного в рамках представленного исследования прототипа экспертной системы ГенДиЭС для диагностики наследственных болезней определяется глубоким анализом различных литературных источников в сочетании с экспертными оценками клинической картины ЛБН и применением алгоритма сопоставительного анализа выдвигаемых интеллектуальной системой гипотез, что позволило добиться всего 10% ошибок на основе данных долабораторного этапа обследования больных.

Заключение

Результат тестирования прототипа экспертной системы для диагностики ЛБН, представленный на примере МПС, продемонстрировал эффективность предложенного подхода к извлечению, представлению и последующей обработке экспертных знаний с помощью построенных моделей и алгоритма сопоставительного анализа интегральных оценок с персональными оценками в конкретном случае.

Участие двух экспертов способствовало объективизации экспертных оценок вследствие формирования коллективного мнения, нивелирующего субъективность присутствующую в знаниях и опыте любого человека.

В перспективе возможно расширение спектра наследственных болезней и создание мультинозологической облачной диагностической системы.

Литература

- Пузырев В.П. Медицинская патогенетика. *Вавиловский журнал генетики и селекции*. 2014;18(1):7–21.
- Новиков П.В., Семьякина А.Н. Лизосомные болезни накопления. В кн.: Основы внутренней медицины; под ред. В.С. Моисеева. М: ГЭОТАР-Медиа; 2014:780–792.
- Scarpa M., Harmatz P.R., Meesen B., Giugliani R. Outcomes of a physician survey on the type, progression, assessment, and treatment of neurological disease in Mucopolysaccharidoses. *Journal of Inborn Errors of Metabolism and Screening*. 2018;6:170022. DOI: 10.1177/2326409818759370.
- Назаренко Л.П., Назаренко М.С. Особенности раннего проявления лизосомных болезней накопления. *Медицинская генетика*. 2013; 12(9):20–24.
- Platt F.M., d'Azzo A., Davidson B.L., Neufeld E.F., Tiff C.J. Lysosomal storage diseases. *Nat. Rev. Dis. Primers*. 2018;4(1):27. DOI: 10.1038/s41572-018-0025-4.
- Kuiper G.-A., Meijer O.L.M., Langereis E.J., Wijburg F.A. Failure to shorten the diagnostic delay in two ultra-orphan diseases (mucopolysaccharidosis types I and III): Potential causes and implications. *Orphanet J. Rare Dis*. 2018;13(1):2. DOI: 10.1186/s13023-017-0733-y.
- Parini R., Andria G. Lysosomal storage diseases: Early diagnosis and new treatments. Montrouge: John Libbey Eurotext; 2010:192.
- Blöß S., Klemann C., Rother A.K., Mehmecke S., Schumacher U., Mücke U. et al. Diagnostic needs for rare diseases and shared prediagnostic phenomena: Results of a German-wide expert Delphi survey. *PLoS One*. 2017;12(2):e0172532. DOI: 10.1371/journal.pone.0172532.
- 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(1):69. DOI: 10.1186/s13023-019-1040-6.
- Alves R., Piñol M., Vilaplana J., Teixidó I., Cruz J., Comas J. et al. Computer-assisted initial diagnosis of rare diseases. *Peer J*. 2016;4:e2211. DOI: 10.7717/peerj.2211.
- Baraitser M., Winter R.M. London dysmorphology database, London Neurogenetics Database & Dysmorphology Photo Library on CD-ROM; by M. Baraitser and R.M. Winte; 3rd ed. Oxford: Oxford University Press; 2001.
- Allanson J.E., Cunniff C., Hoyme H.E., McGaughan J., Muenke M., Neri G. Elements of morphology: standard terminology for the head and face. *Am. J. Med. Gen. A*. 2009;149A(1):6–28. DOI: 10.1002/ajmg.a.32612.
- Fryer A. POSSUM (Pictures of Standard Syndromes and Undiagnosed Malformations). *Journal of Medical Genetics*. 1991;28(1):66–67.
- Гаврилова Т.А. Извлечение знаний: лингвистический аспект. *Intelligent enterprise: RE (Корпоративные системы)*. 2001;27(10): 24–28.
- Гаврилова Т.А., Кудрявцев Д.В., Муромцев Д.И. Инженерия знаний. Модели и методы: учебник. СПб.: Издательство «Лань»; 2016:324.
- Кобринский Б.А. Триединство факторов уверенности в задачах медицинской диагностики. *Искусственный интеллект и принятие решений*. 2018;(2):62–72. DOI: 10.14357/20718594180205.
- Грибова В.В., Петряева М.В., Окунь Д.Б., Шалфеева Е.А. Онтология медицинской диагностики для интеллектуальных систем поддержки принятия решений. *Онтология проектирования*. 2018;8(1):58–73. DOI: 10.18287/2223-9537-2018-8-1-58-73.
- Воинов А.В., Кобринский Б.А., Демикова Н.С. Интеллектуальный анализ медицинских данных с использованием процедуры словарного шкалирования. Двенадцатая национальная конференция по искусственному интеллекту с международным участием КИИ-2010: труды конференции. Т. 1. М.: Физматлит; 2010:153–160.
- Kobrinskii B.A., Blagosklonov N.A. Hybrid approach to knowledge extraction: Textual analysis and evaluations of experts. *Open Semantic Technologies for Intelligent Systems*. 2018;2(8):191–195.
- Köhler S., Schulz M.H., Krawitz P., Bauer S., Dölken S., Ott C.E. et al. Clinical diagnostics in human genetics with semantic similarity searches in ontologies. *Am. J. Hum. Genet*. 2009;85(4):457–464. DOI: 10.1016/j.ajhg.2009.09.003.
- Семьякина А.Н., Новиков П.В., Воскобоева Е.Ю., Захарова Е.Ю., Букина Т.М., Смирнова Г.В. и др. Мукополисахаридозы у детей. *Российский вестник перинатологии и педиатрии*. 2007;52(4): 22–29.
- by M. Baraitser and R.M. Winte; 3rd ed. Oxford: Oxford University Press; 2001.
- Allanson J.E., Cunniff C., Hoyme H.E., McGaughan J., Muenke M., Neri G. Elements of morphology: standard terminology for the head and face. *Am. J. Med. Gen. A*. 2009;149A(1):6–28. DOI: 10.1002/ajmg.a.32612.
- Fryer A. POSSUM (Pictures of Standard Syndromes and Undiagnosed Malformations). *Journal of Medical Genetics*. 1991;28(1):66–67.
- Гаврилова Т.А. Knowledge extraction: a linguistic aspect. *Intelligent Enterprise: RE (Corporate Systems)*. 2001;27(10):24–28 (In Russ.).
- Gavrilova T.A., Kudryavcev D.V., Muromcev D.I. Knowledge engineering. Models and methods: Textbook. St.-Petersburg: Izdatel'stvo "Lan"; 2016:324 (In Russ.).
- Kobrinskii B.A. Certainty factors triunity in the medical diagnostics tasks. *Artificial Intelligence and Decision Making*. 2018;(2):62–72 (In Russ.). DOI: 10.14357/20718594180205.
- Gribova V.V., Petryaeva M.V., Okun D.B., Shalfeeva E.A. Medical diagnosis ontology for intelligent decision support systems. *Ontology of Designing*. 2018;8(1):58–73 (In Russ.). DOI: 10.18287/2223-9537-2018-8-1-58-73.
- Voinov A.V., Kobrinskii B.A., Demikova N.S. Mining medical data analysis using the dictionary scaling procedure. Twelfth National Conference on Artificial Intelligence with International Participation CAI-2010: Proc. V. 1. Moscow: Fizmatlit; 2010:153–160 (In Russ.).
- Kobrinskii B.A., Blagosklonov N.A. Hybrid approach to knowledge extraction: Textual analysis and evaluations of experts. *Open Semantic Technologies for Intelligent Systems*. 2018;2(8):191–195.
- Köhler S., Schulz M.H., Krawitz P., Bauer S., Dölken S., Ott C.E. et al. Clinical diagnostics in human genetics with semantic similarity searches in ontologies. *Am. J. Hum. Genet*. 2009;85(4):457–464. DOI: 10.1016/j.ajhg.2009.09.003.
- Semyachkina A.N., Novikov P.V., Voskoboyeva E.Yu., Zakharova E.Yu., Bukina T.M., Smirnova G.V. et al. Mucopolysaccharidoses in children. *Russian Bulletin of Perinatology and Pediatrics*. 2007;52(4):22–29 (In Russ.).

References

- Puzyrev V.P. Medical pathogenetics. *Vavilov Journal of Genetics and Breeding*. 2014;18(1):7–21 (In Russ.).
- Novikov P.V., Semyachkina A.N. Lysosomal diseases. In: Internal medicine basis; Moiseev V.S. (ed.). Moscow: GEOTAR-Media; 2014:780–792 (In Russ.).
- Scarpa M., Harmatz P.R., Meesen B., Giugliani R. Outcomes of a physician survey on the type, progression, assessment, and treatment of neurological disease in Mucopolysaccharidoses. *Journal of Inborn Errors of Metabolism and Screening*. 2018;6:170022. DOI: 10.1177/2326409818759370.
- Nazarenko L.P., Nazarenko M.S. The early symptoms of lysosomal storage diseases. *Medical Genetics*. 2013;12(9):20–24 (In Russ.).
- Platt F.M., d'Azzo A., Davidson B.L., Neufeld E.F., Tiff C.J. Lysosomal storage diseases. *Nat. Rev. Dis. Primers*. 2018;4(1):27. DOI: 10.1038/s41572-018-0025-4.
- Kuiper G.-A., Meijer O.L.M., Langereis E.J., Wijburg F.A. Failure to shorten the diagnostic delay in two ultra-orphan diseases (mucopolysaccharidosis types I and III): Potential causes and implications. *Orphanet J. Rare Dis*. 2018;13(1):2. DOI: 10.1186/s13023-017-0733-y.
- Parini R., Andria G. Lysosomal storage diseases: Early diagnosis and new treatments. Montrouge: John Libbey Eurotext; 2010:192.
- Blöß S., Klemann C., Rother A.K., Mehmecke S., Schumacher U., Mücke U. et al. Diagnostic needs for rare diseases and shared prediagnostic phenomena: Results of a German-wide expert Delphi survey. *PLoS One*. 2017;12(2):e0172532. DOI: 10.1371/journal.pone.0172532.
- 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(1):69. DOI: 10.1186/s13023-019-1040-6.
- Alves R., Piñol M., Vilaplana J., Teixidó I., Cruz J., Comas J. et al. Computer-assisted initial diagnosis of rare diseases. *Peer J*. 2016;4:e2211. DOI: 10.7717/peerj.2211.
- Baraitser M., Winter R.M. London dysmorphology database, London Neurogenetics Database & Dysmorphology Photo Library on CD-ROM; by M. Baraitser and R.M. Winte; 3rd ed. Oxford: Oxford University Press; 2001.
- Allanson J.E., Cunniff C., Hoyme H.E., McGaughan J., Muenke M., Neri G. Elements of morphology: standard terminology for the head and face. *Am. J. Med. Gen. A*. 2009;149A(1):6–28. DOI: 10.1002/ajmg.a.32612.
- Fryer A. POSSUM (Pictures of Standard Syndromes and Undiagnosed Malformations). *Journal of Medical Genetics*. 1991;28(1):66–67.
- Gavrilova T.A. Knowledge extraction: a linguistic aspect. *Intelligent Enterprise: RE (Corporate Systems)*. 2001;27(10):24–28 (In Russ.).
- Gavrilova T.A., Kudryavcev D.V., Muromcev D.I. Knowledge engineering. Models and methods: Textbook. St.-Petersburg: Izdatel'stvo "Lan"; 2016:324 (In Russ.).
- Kobrinskii B.A. Certainty factors triunity in the medical diagnostics tasks. *Artificial Intelligence and Decision Making*. 2018;(2):62–72 (In Russ.). DOI: 10.14357/20718594180205.
- Gribova V.V., Petryaeva M.V., Okun D.B., Shalfeeva E.A. Medical diagnosis ontology for intelligent decision support systems. *Ontology of Designing*. 2018;8(1):58–73 (In Russ.). DOI: 10.18287/2223-9537-2018-8-1-58-73.
- Voinov A.V., Kobrinskii B.A., Demikova N.S. Mining medical data analysis using the dictionary scaling procedure. Twelfth National Conference on Artificial Intelligence with International Participation CAI-2010: Proc. V. 1. Moscow: Fizmatlit; 2010:153–160 (In Russ.).
- Kobrinskii B.A., Blagosklonov N.A. Hybrid approach to knowledge extraction: Textual analysis and evaluations of experts. *Open Semantic Technologies for Intelligent Systems*. 2018;2(8):191–195.
- Köhler S., Schulz M.H., Krawitz P., Bauer S., Dölken S., Ott C.E. et al. Clinical diagnostics in human genetics with semantic similarity searches in ontologies. *Am. J. Hum. Genet*. 2009;85(4):457–464. DOI: 10.1016/j.ajhg.2009.09.003.
- Semyachkina A.N., Novikov P.V., Voskoboyeva E.Yu., Zakharova E.Yu., Bukina T.M., Smirnova G.V. et al. Mucopolysaccharidoses in children. *Russian Bulletin of Perinatology and Pediatrics*. 2007;52(4):22–29 (In Russ.).

Информация о вкладе авторов

Кобринский Б.А. предложил концепцию исследования, осуществил сравнение с аналогичными системами, провел экспертную оценку коэффициентов модальности и факторов уверенности в качестве одного из экспертов при участии Благосклонова Н.А. как когнитолога.

Благосклонов Н.А. разработал шаблон текстологической карты, провел анализ литературных источников на русском и английском языках с последующим заполнением соответствующих текстологических карт, осуществил перенос извлеченных экспертных знаний в онтологическую систему IACPaaS, занимался разработкой и тестированием алгоритмов по дифференциальной диагностике наследственных болезней обмена.

Благосклонов Н.А. и Кобринский Б.А. анализировали полученные результаты и давали им интерпретацию.

Оба автора внесли существенный вклад в подготовку публикации, дали окончательное согласие на подачу рукописи и согласились нести ответственность за все аспекты работы, ручаясь за их точность и безупречность.

Сведения об авторах

Благосклонов Николай Алексеевич, инженер-исследователь отдела систем поддержки принятия клинических решений, Федеральный исследовательский центр «Информатика и управление» Российской академии наук. ORCID 0000-0002-5293-8469.

E-mail: nblagosklonov@gmail.com.

Кобринский Борис Аркадьевич, д-р мед. наук, профессор, ведущий отделом систем поддержки принятия клинических решений, Федеральный исследовательский центр «Информатика и управление» Российской академии наук. ORCID 0000-0002-3459-8851.

E-mail: kba_05@mail.ru.

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

Information on author contributions

Kobriniskii B.A. proposed a research concept, carried out a comparison with similar systems. Blagosklonov N.A. developed a template for textological cards, carried out an analysis of literary sources in Russian and English, followed by filling in the corresponding textological cards. Kobriniskii B.A., as one of the experts, carried out an expert assessment of modality coefficients and certainty factors, with participation of N.A. Blagosklonov, as a cognitologist. The transfer of the extracted expert knowledge to the IACPaaS ontological system was carried out by N.A. Blagosklonov. N.A. Blagosklonov was developer and tester of algorithms for the differential diagnosis of hereditary metabolic diseases. B.A. Kobriniskii together with N.A. Blagosklonov analyzed the results obtained and gave them an interpretation.

Both authors made significant contributions to the preparation of the publication and gave their final consent to the submission of the manuscript and agreed to be responsible for all aspects of the work, vouching for their accuracy and flawlessness.

Information about the authors

Nikolay A. Blagosklonov, Engineer-Researcher, Department of Decision Clinical Support Systems, Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences. ORCID 0000-0002-5293-8469.

E-mail: nblagosklonov@gmail.com.

Boris A. Kobriniskii, Dr. Sci. (Med.), Professor, Head of the Department of Decision Clinical Support Systems, Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences. ORCID 0000-0002-3459-8851.

E-mail: kba_05@mail.ru.

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

Поступила 04.08.2020

Received August 04, 2020