

ПРИМЕНЕНИЕ ЦИФРОВЫХ ТЕХНОЛОГИЙ В РОССИЙСКИХ УЧРЕЖДЕНИЯХ ЗДРАВООХРАНЕНИЯ ДЛЯ ВЫЯВЛЕНИЯ ПЕРВИЧНОГО ИММУНОДЕФИЦИТА У ДЕТЕЙ¹

THE USE OF DIGITAL TECHNOLOGIES IN RUSSIAN HEALTHCARE INSTITUTIONS TO DETECT PRIMARY IMMUNODEFICIENCY IN CHILDREN

**E. Zaitseva
L. Voronina
S. Deryabina**

Summary. The article examines the general state of the use of digital technologies in Russian healthcare institutions, the use of digital technologies in molecular genetic research, including for the detection of primary immunodeficiency in children. The authors prove that the use of digital technologies for diagnostics and timely molecular genetic verification of the diagnosis is of applied practical importance, since allows to accelerate the planning and implementation of radical therapy methods for the complete cure of patients with primary immunodeficiency. Also, carrying out molecular genetic studies using digital technologies in patients with primary immunodeficiency allows discovering new variants of nucleotide changes in DNA, accumulating data on the genetic heterogeneity of primary forms of immunodeficiency, and expanding general knowledge in the field of immunogenetics.

Keywords: digital technologies, primary immunodeficiency, molecular genetic research, patients.

Зайцева Екатерина Васильевна

*К.с.н., доцент, Уральский федеральный университет
(Екатеринбург)
e.v.zaitceva@urfu.ru*

Воронина Людмила Ивановна

*К.с.н., доцент, Уральский федеральный университет
(Екатеринбург)
voronina_612@outlook.com*

Дерябина Светлана Степановна

*К.б.н., Клинико-диагностический центр «Охрана
здоровья матери и ребенка» (Екатеринбург); н.с.,
Институт иммунологии и физиологии УрО РАН; н.с.,
Уральский федеральный университет (Екатеринбург)
ssderyabina@gmail.com*

Аннотация. В статье рассматривается общее состояние применения цифровых технологий в российских учреждениях здравоохранения, в молекулярно-генетических исследованиях, и в том числе для выявления первичного иммунодефицита у детей. Авторы доказывают, что подобный подход имеет прикладное практическое значение для диагностики и своевременной молекулярно-генетической верификации диагноза, т.к. позволяет ускорить планирование и реализацию радикальных методов терапии для полного излечения пациентов с первичными иммунодефицитами. Проведение молекулярно-генетических исследований с применением цифровых технологий у пациентов с первичным иммунодефицитом позволяет также открывать новые варианты нуклеотидных изменений ДНК, накапливать данные о генетической гетерогенности первичных форм иммунодефицитов, расширять общие знания в области иммуногенетики.

Ключевые слова: цифровые технологии, первичный иммунодефицит, молекулярно-генетические исследования, пациенты.

Тема применения цифровых технологий (искусственного интеллекта, Big Data и др.) в здравоохранении активно обсуждается как теоретиками, так и практиками. Все сходятся во мнении, что благодаря этим технологиям повышается качество медицинского обслуживания, повышается точность диагностики. Дискуссии развиваются по нескольким направлениям: применение искусственного интеллекта как инструмента поддержки в принятии решений по диагности-

ке различных заболеваний (Ш.Х. Ганцев, М.В. Франц); эффективность использования такой разновидности технологий искусственного интеллекта, как консультативные интеллектуальные медицинские системы (Б.А. Кобринский); развитие технологий искусственного интеллекта будущими врачами (К. Paranjare, M. Schinkel, R. Nannan, P. Nanayakkara); совершенствование технологий «BIG DATA» в медицине (Н.С. Карнаухова, Р.Г. Ильюхин), в том числе технологий секвенирования нового

¹ Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 18–29–14059 «Правовые и этические аспекты технологий геномных исследований при врожденных ошибках иммунитета».

поколения (A. Tarasov, A. J. Vilella, E. Cuppen, I. J. Nijman, P. Prins). В то же время весьма мало исследований, посвященных использованию цифровых технологий в молекулярно-генетических исследованиях, проводимых при диагностике первичных иммунодефицитов у детей (И. А. Тузанкина, С. С. Дерябина, М. А. Болков). Данное обстоятельство подтверждает актуальность темы, заявленной в настоящей статье. Цель исследования, результаты которого представлены в статье,— изучение применения цифровых технологий в отечественных учреждениях здравоохранения.

Россия активно осваивает цифровые технологии и внедряет их в различные сферы деятельности, включая здравоохранение. В публичной декларации приоритетных направлений, заявленных на 2020 г. Министерством здравоохранения РФ, предполагается применять их для совершенствования управленческой деятельности путем обмена электронными документами на базе Единой государственной информационной системы, подключить медицинские организации к подсистемам централизованных государственных информационных систем (см. Публичная декларация приоритетных направлений деятельности министерства здравоохранения Российской Федерации на 2020 год). Ожидается, что применение этих новейших технологий позволит сделать более совершенным статистический учет заболеваемости и смертности, повысить оперативность в принятии управленческих решений на основе достоверных клинических и лабораторных данных, улучшить качество и доступность медицинской помощи, прежде всего высокотехнологичной.

Понятие «цифровые технологии в здравоохранении», безусловно, собирательное. В данном случае предполагается, что используются технологии, основанные на искусственном интеллекте, к возможностям которого относятся распознавание изображений, речи, перевод с одного языка на другой, выбор диагностического направления по имеющимся алгоритмам или решение задач, для которых не существует алгоритмов, но есть гарантия получения достоверных результатов. Безусловно, что возможности применения этих технологий активно изучаются не только для управления в здравоохранении, но и для диагностики некоторых заболеваний [2], для определения тактики лечения, например, в инвазивной хирургии [11].

Искусственный интеллект применяется в консультативных интеллектуальных системах, что позволяет диагностировать множество нозологических единиц. Так, база знаний ЭС INTERNIST-I/CADUCEUS (наукоемкая медицинская экспертная система) [8] содержит информацию о более чем 500 терапевтических заболеваниях [4]. В интернете появляются публикации о том, что между-

народные компании активно продвигают свои проекты по использованию искусственного интеллекта в диагностике. Так, Google планирует собирать сотни различных образцов генетического материала добровольцев, проанализировать информацию с помощью вычислительных мощностей своих компьютерных систем для выявления биомаркеров и индикаторов определенного заболевания (Baseline — новый проект Google по созданию генетически-молекулярной карты полностью здорового человека).

Версия Deep Variant, предназначенная для анализа генетических данных, пока считается самой точной, позволяющей назначать наиболее эффективное лечение [11]. В профессиональном сообществе клинических генетиков большую популярность имеет приложение Face2Gene — сертифицированная программа, основанная на «глубинном обучении» и «искусственном интеллекте», при помощи которой загруженные фото пациентов сравниваются с собранными в единую базу данных сотен лиц пациентов с редкими генетическими болезнями для обнаружения ключевых дисморфологических особенностей, поиска ассоциаций «фенотип-генотип» и помощи в постановке диагноза. Приложение позволяет осуществлять совместное рассмотрение кейсов медицинскими специалистами и выработки оптимальных алгоритмов для решения возникающих диагностических дилемм. Разработчики, однако, подчеркивают, что данное ПО является только инструментом для поиска нужной информации и выбора направления исследования, и не предназначено для замены очного медико-генетического консультирования, и тем более, выполнения самодиагностики и назначения самолечения.

Также к цифровым технологиям, используемым в здравоохранении, относятся технологии Big Data. Впрочем, это понятие тоже собирательное и охватывает множество подходов и технологий. Особые надежды практики возлагают на применение технологий Big Data для анализа молекулярных данных, получаемых путем изучения структуры молекул ДНК, молекул РНК, белковых молекул, взаимодействия клеточных метаболитов. Ожидаемый результат — это скорость генерации информации, ее обработки, эффективного анализа полученных данных. Возможность обработки больших данных позволила разработать технологии секвенирования нового поколения [13], которые называют революционными из-за высокой скорости секвенирования генома и транскриптома, а также получения высокой экономической эффективности и производительности труда специалистов клинической и лабораторной генетики [12].

Современные подходы в анализе и интерпретации данных молекулярно-генетических исследований в диагностике врожденных ошибок иммунитета также пред-

полагают использование технологии искусственного интеллекта. В частности, появившаяся несколько лет назад и уверенно занявшая передовые позиции в данной области технология NGS (next generation sequences) столкнулась не только с техническими проблемами, но и с проблемой хранения огромных объемов получаемых сырых данных и файлов FASTQ (кодированная текстовая информация, описывающая биологическую последовательность) и нехваткой вычислительных мощностей для их корректного и быстрого анализа, а также со сложностью в интерпретации полученных данных для принятия врачебного решения [1]. Патогенные и вероятно патогенные варианты в основном уже отражены в существующих международных базах данных: Online Mendelian Inheritance in Man (OMIM), The Human Gene Mutation Database (HGMD), ClinVar, The Genome Aggregation Database (gnomAD). Однако необходимо понимать, что генетика и молекулярная биология являются сегодня едва ли не самыми быстроразвивающимися областями науки, и обновление их научных данных происходит практически ежедневно, чего нельзя сказать о бесплатных электронных ресурсах, связанных с ними, которые добавляют или изменяют свои данные не чаще одного раза в два–три года. Доступ же к платным базам этих зарубежных компаний (обновляемым на постоянной основе), как правило, отсутствует в большинстве российских клиник и институтов.

Проблема усугубляется тем, что имеющиеся международные базы данных слабо применимы к некоторым российским популяциям, а собственная БД по генетическим вариантам существует в виде фрагментарных и разрозненных результатов исследований отдельных научных институтов. Кроме того, зачастую сырые данные демонстрируют генетические варианты, которые ранее не были описаны в мире и выявлены у пациента впервые. Как определить, патогенны они для организма или безвредны? Ключевым критерием для клинической интерпретации патогенности-доброкачественности рассматривают частоту варианта в общей популяции: при частоте аллели выше ожидаемой для заболевания аргумент считается в пользу не патогенности варианта [5]. В таких случаях возможно проведение экспериментов по функциональному анализу, и существуют специальные лаборатории функциональной геномики, использующие современные подходы для исследования белок-белковых взаимодействий, изучения локализации белка и анализа сплайсинга на модельных животных или культурах клеток. Для каждого такого варианта разрабатывается свой уникальный дизайн эксперимента, помогающий определить патогенность исследуемого изменения ДНК. Однако такие лаборатории, как правило, единичные, и не в состоянии справиться с всевозрастающим потоком «генетического бума». И тогда на помощь специалистам приходят биоинформатические инстру-

менты: Alamut, Mutation Taster, Polyphen2, SIFT и др. [7, 10]. Это своего рода программы-предикторы, на основании математических расчетов предсказывающие эффект, который может оказать исследуемый генетический вариант на организм человека.

Подобного рода находки могут встречаться даже при использовании традиционных методов молекулярно-генетической диагностики. Так, в лаборатории молекулярной диагностики КДЦ «Охрана здоровья матери и ребенка» (г. Екатеринбург) отсутствует технология NGS, и поиск генетических причин развития иммунозависимой патологии проводится традиционными доступными методами: мультиплексная лигазозависимая амплификация проб (MLPA), технология VACs-on-Beads, ПЦР (полимеразная цепная реакция) в режиме реального времени, таргетное секвенирование генов-кандидатов, однако обнаружение новых, неописанных ранее вариантов, за последние три года перестало быть редкостью. Особенно богатым на подобный «улов» можно назвать заболевание из группы ПИД-Х-сцепленную агаммаглобулинемию. Всего в гене брутонтирозинкиназы (ВТК), ассоциированном с расстройством гуморального иммунитета вследствие полного или частичного отсутствия В-клеток, известно более 500 различных вариантов, включающих единичные замены нуклеотидов, дефекты сайтов сплайсинга, а также короткие делеции и инсерции.

В зарубежных исследованиях нередко рассматриваются вопросы прогнозирования клинических проявлений в отношении корреляции генотип-фенотип с указанием того, что некоторые из мутаций могут вызывать более высокую степень нарушения гуморального иммунитета и раннюю манифестацию заболевания, однако интерпретация ранее неописанных вариантов и прогноз их влияния на здоровье пациента все еще вызывают затруднения. За годы работы 11 пациентов с дефицитом В-клеточного звена были направлены иммунологами Екатеринбурга и Челябинска на исследование по поводу Х-сцепленной агаммаглобулинемии. Молекулярно-генетическое подтверждение диагноза получили семь детей. Выявленные в лаборатории методом таргетного секвенирования кодирующей части гена варианты, приведшие к развитию патологии, включали в себя как делеции (от участков в несколько нуклеотидов до протяженных, включающих целые экзоны), так и миссенс-нонсенс мутации. При этом почти половина находок — 3 варианта из детектированных 7 (делеция 13-ти нуклеотидов во втором экзоне [9] и две инсерции аденина в 929 и 1052 положении комплементарной ДНК — кДНК) — совершенно новые, не описанные ранее изменения нуклеотидной последовательности гена ВТК. Характер наследования и алгоритмы предсказания патогенности данных вариантов (Mutation Taster,

PolyPhen2.0., SIFT) позволяют сделать выводы об их каузативном значении.

Несомненно, в отсутствие экспериментальных данных по изучению трансгенных организмов с изучаемой мутацией существующие базы данных и программные продукты полезны для более полного анализа данных, уточнения свойств транскрипта, местоположения гена и предсказания патогенности, однако автоматическое использование только предикторов генетического эффекта может приводить как к ложноположительным, так и к ложноотрицательным результатам. Важно помнить, что только тщательно собранный семейный анамнез, строгие критерии отбора вариантов, сопоставление каждого из них с описанием фенотипа пациента, использование достаточно крупной выборки позволят сделать обоснованные выводы о патогенности неизвестных мутаций.

И все же применение цифровых технологий позволяет значительно обогащать собственные базы данных и самостоятельно проводить анализ корреляции «генотип-фенотип» при выявлении новых локусов ген-

ных ошибок. Это также способствует более глубокому пониманию патофизиологии болезни и индивидуальному подходу к выбору методов паллиативной или радикальной терапии, и в конечном итоге увеличивает вероятность благоприятного прогноза для каждого конкретного пациента [6]. Таким образом, интегральный показатель результативности применения различных технологий искусственного интеллекта — это, конечно, улучшение качества медицинских услуг. И все же, при всей значимости применения технологий искусственного интеллекта, ведущими остаются такие компетенции специалистов-медиков, как умение анализировать информацию, интерпретировать результаты, взаимодействовать с пациентами. Практика деятельности научной группы под руководством И. А. Тузанкиной еще раз подтверждает вывод о том, что использование «цифрового разума» для ускорения аннотирования генетических данных и анализа генного полиморфизма в диагностике наследственных заболеваний при всей его потенциальной мощности все же отдает главенствующую роль в этой области науки естественному человеческому интеллекту.

ЛИТЕРАТУРА

1. Бархатов И.М., Предеус А. В., Чухловин А. Б. Секвенирование нового поколения и области его применения в онкогематологии // Онкогематология. 2016. № 4. С. 56–63.
2. Ганцев Ш.Х., Франц М. В. Искусственный интеллект как инструмент поддержки в принятии решений по диагностике онкологических заболеваний // Медицинский вестник Башкортостана. 2018. Т. 13. № 4 (76). С. 67–71.
3. Карнаухов Н.С., Ильяхин Р.Г. Возможности технологий «BIG DATA» в медицине // Врач и информационные технологии. 2019. № 1. С. 59–63.
4. Кобринский Б. А. Консультативные интеллектуальные медицинские системы: классификации, принципы построения, эффективность // Врач и информационные технологии. 2008. № 2. С. 38–47.
5. Рыжкова О.П., Кардымон О. Л., Прохорчук Е. Б., Коновалов Ф. А., Масленников А. Б., Степанов В. А., Афанасьев А. А., Захлязьминская Е. В., Костарева А. А., Павлов А. Е., Голубенко М. В., Поляков А. В., Куцев С. И. Руководство по интерпретации данных, полученных методами массового параллельного секвенирования (MPS) // Медицинская генетика. 2017. № 16 (7). С. 4–17.
6. Тузанкина И. А. Дерябина С. С., Болков М. А. Первичные иммунодефициты в раннем возрасте. Москва: РАН, 2018. 176 с.
7. Adzhubei I., Jordan D. M., Sunyaev S. R. Predicting functional effect of human missense mutations using PolyPhen-2 // Current protocols in human genetics. 2013. Т. 76. № 1. P. 7.20.1–7.20.41.
8. Banks G. Artificial intelligence in medical diagnosis: the INTERNIST/CADUCEUS approach // Critical reviews in medical informatics. 1986. Т. 1. № 1. С. 23–54.
9. Deryabina S., Tuzankina, I., Vlasova E., Pavlova A., Bolkov M. A. Novel Bruton's Tyro-sine Kinase Mutation in Russian Patient with X-Linked Agammaglobulinemia // Case Reports in Clinical Medicine. 2017. № 6. P. 241–249.
10. Kumar P., Henikoff S., Ng P. C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm // Nature protocols. 2009. Т. 4. № 7. С. 1073–1081.
11. Paranjape K., Schinkel M., Nannan R., Nanayakkara P. Introducing Artificial Intelligence Training in Medical Education // JMIR Medical Education. 2019. Vol. 5. Issue 2. P. e16048.
12. Tarasov A., Vilella A. J., Cuppen E., Nijman I. J., Prins P. Sambamba: fast processing of NGS alignment formats // Bioinformatics. 2015. Vol. 31. Issue 12. P. 2032–2034.
13. Zaitseva E., Bolkov M., Tuzankina I. Application of genomic research technologies for innate immunity errors // 6th International Multidisciplinary Scientific Conference SOCIAL SCIENCES & ARTS SGEM 2019. Conference Proceeding. Volume 6. Science and Society. Issue 3. Education and Educational research. Albena: STEF92 Technology Ltd., 2019. P. 393–400.

© Зайцева Екатерина Васильевна (e.v.zaitceva@urfu.ru),

Воронина Людмила Ивановна (voronina_612@outlook.com), Дерябина Светлана Степановна (ssderyabina@gmail.com).

Журнал «Современная наука: актуальные проблемы теории и практики»