

ISSN 2782-3806  
ISSN 2782-3814 (Online)  
УДК 61:004.418

## ИНФОРМАЦИЯ КАК ВАЖНЕЙШИЙ ИНСТРУМЕНТ РАЗВИТИЯ ПЕРСОНАЛИЗИРОВАННОЙ МЕДИЦИНЫ. КАК НАУЧИТЬСЯ ЕЙ УПРАВЛЯТЬ НА БЛАГО ПАЦИЕНТА. НАУКА О «БОЛЬШИХ ДАННЫХ»

**Шляхто Е. В., Конради А. О., Курапеев Д. И.**

Федеральное государственное бюджетное учреждение «Национальный медицинский исследовательский центр имени В. А. Алмазова» Министерства здравоохранения Российской Федерации, Санкт-Петербург, Россия

**Контактная информация:**

Конради Александра Олеговна,  
ФГБУ «НМИЦ им. В. А. Алмазова»  
Минздрава России,  
ул. Аккуратова, д. 2, Санкт-Петербург,  
Россия, 197341.  
E-mail: konradi@almazovcentre.ru

Статья поступила в редакцию 03.11.2022  
и принята к печати 21.11.2022.

### РЕЗЮМЕ

Обзор содержит информацию о современном состоянии применения аналитики больших данных в медицине, включая анализ омиксных данных. Обсуждаются вызовы, особенности больших данных в здравоохранении, причины недостаточно быстрого внедрения аналитики больших данных и алгоритмов искусственного интеллекта, источники данных и перспективы применения в ближайшие годы.

**Ключевые слова:** большие данные, омиксные данные, персонализированная медицина, электронное здоровье.

*Для цитирования: Шляхто Е.В., Конради А.О., Курапеев Д.И. Информация как важнейший инструмент развития персонализированной медицины. Как научиться ей управлять на благо пациента. Наука о «больших данных». Российский журнал персонализированной медицины. 2022;2(6):6-15. DOI: 10.18705/2782-3806-2022-2-6-6-15.*

# INFORMATION AS A TOOL FOR PERSONALIZED MEDICINE DEVELOPMENT. HOW TO RULE FOR PATIENT BENEFIT. BIG DATA SCIENCE

Shlyakhto E. V., Konradi A. O., Kurapeev D. I.

Almazov National Medical Research Centre, Saint Petersburg, Russia

**Corresponding author:**

Konradi Aleksandra O.,  
Almazov National Medical Research Center,  
Akkuratova str., 2, Saint Petersburg,  
Russia, 197341.  
E-mail: konradi@almazovcentre.ru

Received 03 November 2022; accepted  
21 November 2022.

## ABSTRACT

The article summarizes modern approaches of big data analytics implementation in healthcare, including OMIC's data. The existing challenges for big data in medicine, barriers for implementation, reasons for slow development of practical usage of big data analytics and artificial intelligence in healthcare are discussed as well as perspective for the nearest future.

**Key words:** big data, e-health, omics, personalized medicine.

*For citation: Shlyakhto EV, Konradi AO, Kurapeev DI. Information as a tool for personalized medicine development. How to rule for patient benefit. Big data science. Russian Journal for Personalized Medicine. 2022;2(6):6-15. (In Russ.) DOI: 10.18705/2782-3806-2022-2-6-6-15.*

## ВВЕДЕНИЕ

XXI век стал веком информационных технологий, и медицина не является исключением в ряду всех происходящих в мире эволюционных процессов. Накопление данных и современные методы их обработки создают беспрецедентные возможности для выхода здравоохранения на принципиально иные результаты и показатели эффективности. Тем не менее, корректный сбор информации, доступ к ней, алгоритмы обработки и все виды анализа, а также практическое внедрение потенциала data science (науки о данных) и технологий искусственного интеллекта (ИИ) являются одновременно и ведущими вызовами современности. Объемы производимых и потенциально применимых для нужд здравоохранения данных настолько велики, что они требуют модернизации всех подходов к хранению, аналитике и использованию этих данных [1].

В частности, одним из таких вызовов является хранение и обработка геномных и других омиксных данных. Объем данных полногеномного секвенирования занимает примерно столько же места, как 100 кинофильмов с самым высоким разрешением или 150 гигабайт информационного хранилища [2]. В результате любое практическое внедрение концепции персонализированной медицины на основе геномных данных требует модернизации информационной инфраструктуры и создания дополнительных сервисов. Кроме того, ситуацию осложняет дефицит специалистов, способных извлекать из этих данных клинически значимую информацию и осуществлять ее корректную интерпретацию.

Информатизация здравоохранения началась достаточно давно, и на сегодняшний день большинство медицинских организаций в мире уже пользуется теми или иными сервисами, в первую очередь медицинскими информационными системами. Однако движение от простого хранения данных и использования минимальных аналитических функций к предсказательному моделированию и системам поддержки принятия решений происходит крайне медленно. Развитие возможностей аналитики больших данных и технологий ИИ превышает по скорости процессы внедрения во много раз. В целом на сегодняшний день медицинские организации и медицинские работники еще не готовы в полной мере к активному применению информационных технологий нового поколения. Основными барьерами к их внедрению являются:

- Отсутствие должных навыков и знаний у медицинского персонала;

- Отсутствие возможности для систематизированного получения, хранения, архивирования и обработки медицинских данных;

- Низкое качество данных и низкая цифровая культура в медицинских организациях;

- Дефицит средств и единой концепции по формированию корректной и современной информационной инфраструктуры учреждений;

- Сложности при защите персональных данных и этические вопросы доступа к ним;

- Консервативность медицины как отрасли и отсутствие готовности организации к полной прозрачности данных [3].

Кроме того, развитие возможности информационных систем и генерация новых знаний происходят настолько быстро, что ни концепции развития, ни медицинские работники, ни бизнес-сообщество не успевают адаптироваться под эту лавину новых знаний и сервисов. Постоянно наблюдается как идеологическое, так и техническое отставание медицины от ряда других отраслей экономики, где внедрение нового происходит быстрее и менее болезненно.

Тем не менее, необходимость такого внедрения и максимально быстрой адаптации медицины к растущим возможностям и новым вызовам очевидны. Это потребует в ближайшее время комплексных изменений, включающих смену инфраструктуры, идеологии управления медицинской помощью, подготовки новых кадров и смены самой структуры распределения рабочего времени медицинских работников, вовлечения новых игроков в получение и обработку данных и реализации новых моделей здравоохранения, таких как пациент-центричность, ценностное здравоохранение и персонализированная медицина [4–5]. При этом очень важно понимать, что данные сами по себе, даже очень «большие» и очень качественные, еще ничего не объясняют и не помогают в работе врача. Важно выстроить не просто сбор данных и их упорядочивание, а суметь определить пути поиска нужных связей и корреляций, сформировать корректные алгоритмы обработки и интерпретации, которые дают конкретную выгоду в клинических или управленческих решениях. Именно эта работа представляет собой самую большую сложность, так как ее результат требует валидации в разных условиях, проверки на практике и интеграции с существующими бизнес-процессами медицинских организаций.

Данные, генерируемые здравоохранением, растут в объеме примерно на 50 % ежегодно, и их общий объем давно превысил 2000 гэксабайт [6]. Если представить себе образно, что все данные о здоровье людей будут представлены в виде столбика монет,

то уже в 2020 году он оказался бы высотой более 82 тысяч километров, что составляет одну треть расстояния до Луны, тогда как еще в 2013 году это было всего лишь 3 % от этого расстояния [7]. Согласно оценкам IBM, в периоды между записями в официальных медицинских картах каждый человек оставляет след более 1 миллиона гигабайт в виде разного рода дополнительных данных, имеющих отношение к здоровью [7]. Эти данные, генерируемые интернетом вещей, социальной активностью, мобильными телефонами и носимыми устройствами, сегодня в принципе мало используются медицинскими организациями, хотя представляют собой неоценимый банк для динамического наблюдения за лицами с факторами риска и хронической патологией [8–9].

### ПОНЯТИЕ «БОЛЬШИХ ДАННЫХ»

Ориентированная на пациента медицинская помощь, или так называемая пациент-центричная модель медицины не может игнорировать постоянное нарастание медицинских данных с точки зрения не только их объема, но и разнообразия, и скорости получения, что настойчиво подталкивает медицину к новой парадигме, которая носит широко употребляемое название «большие данные» (БД, Big data) [10]. Анализ огромного объема, в том числе неоднородных, данных и данных, генерируемых с большой скоростью, что и подразумевает наука о больших данных или аналитика больших данных (АБД), позволяет извлечь максимальную ценность из данных и успешно анализировать связи между различными переменными, описывающие жизненные функции пациента и способные повлиять на его здоровье [11]. Именно эта возможность и является драйвером того, что передовые медицинские организации вкладывают сегодня значительные средства в анализ данных, чтобы облегчить в дальнейшем принятие клинических и управленческих решений [12, 13]. Интеграция данных об уникальных характеристиках человека, клинических симптомах и фенотипе, информация, полученная от диагностической визуализации и при проведении лабораторных анализов, медицинские записи в истории болезни, а также данные генома позволяют развиваться прецизионной медицине для достижения целей прогнозирования исходов и истинной профилактики [14]. Наличие большого количества данных имеет решающее значение сегодня также в интенсивной терапии, в том числе они позволяют быстро ставить диагнозы и предлагать конкретные методы лечения для редких патологий или необычного течения болезни, что невозможно сделать только человеческими ресурсами [15].

В последние годы достаточно быстро распространяются персонализированная диагностика и терапия, в том числе за счет развития АБД и использования интернета вещей, который позволяет собирать разного рода данные, генерируемые огромным количеством устройств, и делать их доступными через интегрированные системы, облегчающие передачу информации [16]. В итоге АБД представляет собой большой набор данных из различных источников, не только здравоохранения, который позволяет проводить более персонализированное лечение, оценивать его эффективность и улучшать процессы в медицинской организации, уменьшая риски, связанные с медицинской помощью, за счет инновационных способов управления и контроля происходящих процессов [17–19].

Современные технологии геномики и постгеномики также позволяют получать огромные объемы необработанных данных о сложных биохимических и регуляторных процессах в живом организме. Эти цифровые данные нередко неоднородны, хранятся в разных форматах. Медицинские данные также чаще всего представлены в разнородных форматах. Они могут быть структурированными, полуструктурированными или неструктурированными; дискретными или непрерывными. Наука о больших данных как раз и занимается такими сложными данными, которые трудно анализировать с помощью традиционного программного или аппаратного обеспечения. АБД охватывает интеграцию разнородных данных, контроль качества данных, их анализ, моделирование, интерпретацию и валидацию [20]. АБД в здравоохранении позволяет анализировать большие наборы данных от тысяч и миллионов пациентов, выявлять кластеры и корреляции между наборами данных, а также разрабатывать прогностические модели с использованием методов интеллектуального анализа данных. АБД в медицине и здравоохранении объединяет компетенции нескольких научных областей, таких как биоинформатика, медицинская визуализация, сенсорная информатика, медицинская информатика и информатизация системы здравоохранения.

Согласно определению Еврокомиссии по регулированию исследований и инноваций в области здравоохранения, под «большими данными» в медицине понимают данные большого объема и разнообразия, включающие биологическую информацию, информацию об образе жизни и окружающей среде, а также клинические данные конкретных индивидуумов и больших когорт в соотношении с их статусом здоровья и благополучия в различные временные точки [21]. Очень важно понимать, что понятию АБД соответствует не число пара-

метров или число испытуемых, а свойства самих данных, которые характеризуют не только объем, но и сложность формата, что не позволяет для их анализа использовать стандартные подходы.

АБД характеризуется целым рядом важных свойств, важных для понимания их сути, которые в литературе принято называть 3Vs [22]:

- **Volume** — объем: количество данных, генерируемых каждую секунду;
- **Variety** — разнообразие: генерируются, накапливаются и используются различные типы данных, даже неструктурированные или полуструктурированные;
- **Velocity** — скорость: относится к генерации данных (которая всегда увеличивается с течением времени).

К этим первым 3V позже были добавлены еще 3V [23]:

- **Veracity** — правдивость или неопределенность данных;
- **Value** — ценность: технологии АБД повышают ценность данных, превращая их в действительно полезную информацию;
- **Variability** — вариабельность: данные по одной и той же теме могут иметь различия, связанные с их форматом или способом получения, и это часто является большой проблемой обработки (различные лабораторные единицы измерения, методики расчета).

В последующем добавились еще 2V [24]:

- **Visualization** — возможность визуализации данных в виде сводных графиков и схем;
- **Virality** — жизнеспособность (наличие смысла и пригодность для анализа).

Недавно была введена дополнительная характеристика АБД [25]:

- **Complexity** — сложность: чем больше размер набора данных, а также их разнообразие и неопределенность, тем выше сложность этих данных для анализа.

В современном мире до 90 % медицинских данных являются неструктурированными, что само по себе определяет сложность их обработки.

Конечно, наиболее ценными и надежными источниками для получения АБД в здравоохранении являются медицинские записи (например, электронные медицинские карты, системы поддержки принятия клинических решений, биомедицинские данные и т. д.) [26]. Велика также ценность внешних источников (лаборатории, аптеки, биометрические и другие данные, полученные непосредственно от пациентов, и др.). Дополнительные источники данных становятся все более доступными, такие как данные, полученные в результате использования

интернета (социальных сетей) и интеллектуальных приложений [27–28]. Для управления и обработки этих данных чаще всего используются облачные вычисления. Разнородные данные при правильной интеграции с медицинскими данными позволяют осуществлять мониторинг состояния здоровья пациентов в различных условиях (больницы, дома престарелых, на дому, в условиях реабилитации и даже в поездках) [29–30]. Этот аспект использования БД считается наиболее важным, поскольку основные ошибки, которые могут привести к ошибочному диагнозу и летальному исходу, происходят вследствие отсутствия мониторинга жизненно важных параметров, а также несоблюдения режима лечения [31]. Следует сказать, что использование облачных вычислений и других инструментов и методов АБД сталкивается с рядом трудностей, связанных с сетевыми сбоями, проблемами безопасности и конфиденциальности данных пациентов [32]. Все это в сочетании с консервативностью здравоохранения как отрасли существенно тормозит практическое использование АБД в медицине. Но, тем не менее, примеры успешного внедрения нарастают огромными темпами.

Вторым огромным направлением АБД в медицине является генерация и анализ так называемых омиксных данных [33]. Омиксные данные — большие массивы данных, собранные на разных уровнях биологических процессов. Термин происходит от английского Omics и включает в себя целый ряд новых биологических дисциплин, характеризующихся общим подходом к методологии изучения биообъектов, а именно: изучение не отдельных молекул и биохимических путей, а всего комплекса данных. Омиксные данные — это совокупность всех молекул определенного уровня, отражающая состояние организма или его части. Омиксные технологии используют высокопроизводительные методы анализа, что ведет за собой генерацию больших массивов данных, анализом которых тоже занимаются data scientists. Это породило создание новой дисциплины на стыке биологии, статистики и компьютерной науки — биоинформатики. Следующим этапом стало применение методов машинного обучения для решения биологических задач, что активно развивается в настоящее время.

Практическое использование омиксных данных сегодня уже достаточно велико. Ведущим направлением является новая идеология создания лекарственных препаратов. Анализ омиксных данных, в том числе с применением ИИ, позволяет выявить новые механизмы развития заболевания и выбрать мишени без заранее сформированной гипотезы о тех сигнальных путях, которые задействованы в развитии забо-

**Таблица 1. Основные понятия омиксных данных и их характеристика (адаптировано и дополнено из [33])**

Наименование	Краткое описание
Геномика	Все данные, характеризующие геном. В настоящее время предметом изучения в основном являются некодирующие последовательности и их роль.
Эпигеномика	Все эпигеномные модификации генома в ядре и в клетке. Регуляция экспрессии генов.
Транскриптомика	Оценка уровня экспрессии генов в конкретной клетке или клеточной популяции.
Протеомика	Оценка всех возможных взаимодействий, которые осуществляются между белками, весь набор белков, который кодируется геномом в конкретном типе клеток и на уровне всего организма, в том числе в зависимости от конкретных условий и конкретного периода времени.
Метаболомика	Изучение спектра метаболитов (малых молекул) в клетке, в органеллах, в органе или на уровне всего организма.
Интерактомика	Изучение прямых и косвенных взаимодействий между белками и иными молекулами внутри конкретной клетки и оценка последствий таких взаимодействий. Описание этих процессов в виде биологических сетей.
Фармакогеномика	Комбинация геномики и фармакологии с анализом роли генома в индивидуальном ответе на лекарство.
Феномика	Описание фенотипа с его бесчисленными количественными и качественными характеристиками.
Болезнемика (diseasomics)	Описание всех существующих болезней и их характеристик, чаще всего в контексте генетических причин.
Микробиомика	Данные о генотипе и фенотипе населяющих организм человека микроорганизмов и их взаимодействиях.

левания. Иными словами, сравнение экспрессии генов в норме и патологии дает основания, например, определять мишени воздействия, не понимая заранее их функциональную значимость. Далее биоинформатика помогает провести валидацию мишеней, определить точный механизм действия кандидата, а также вычислить его токсичность, что в конечном итоге позволяет сделать drug-дизайн более быстрым и персонализированным [34].

## ЭЛЕКТРОННОЕ ЗДРАВООХРАНЕНИЕ

Сегодня медицинские и научные исследования в области медицины уже проводятся не только с использованием традиционных устройств, но и, например, с помощью так называемых интеллектуальных устройств, которые все больше становятся

неотъемлемыми элементами повседневной жизни [35–36]. Продукт технологической революции, которая началась до 2000 года с резкого скачка значимости интернета, а позже и с огромным распространением подключенных к нему устройств нового поколения (IoT), — это и есть электронное здравоохранение. На самом деле электронное здравоохранение обладает огромным потенциалом для повышения эффективности всей системы здравоохранения (сокращение расходов) и эффективности самого лечебного процесса (качество медицинской помощи) [37–39]. Электронное здравоохранение представляет собой основу так называемого Здравоохранения 4.0 [40]. Именно применение новых интеллектуальных технологий, интернета вещей [37], обмена данными между различными областями и системами, робототехники и облачных вы-

числений [41] может привести к улучшению оказания медицинской помощи в будущем.

Развитие электронного здравоохранения идет уже достаточно давно и включает в себя ряд стратегически важных сервисов и организационных решений. Утверждение новых технологий определило создание стандартов цифровой обработки изображений и стандартов передачи данных в медицине (DICOM), определяющих правила хранения и обмен изображениями, выходящий за рамки старого поколения аналоговых машин. Следующим важным рубежом стало создание методологии и правил ведения электронных медицинских карт (EHR), которые через несколько лет полностью заменят бумажные носители. На этой основе сформировались в целом требования и стандарты для медицинских информационных систем. В РФ также утверждены приказом Минздрава Требования к государственным информационным системам в сфере здравоохранения субъектов Российской Федерации, медицинским информационным системам медицинских организаций и информационным системам фармацевтических организаций (<https://base.garant.ru/72217630/>). Более того, последовательно принято 6 частей в Национальный стандарт РФ ГОСТ Р 59921.1-2022 «Системы искусственного интеллекта в клинической медицине», которые регламентируют процессы тестирования и внедрения подобных систем в здравоохранение. Однако до сих пор не решены вопросы применения в реальном здравоохранении обособленного ИИ.

## СУЩЕСТВУЮЩИЕ ПРОБЛЕМЫ АНАЛИТИКИ БОЛЬШИХ ДАННЫХ В МЕДИЦИНЕ

Одной из основных технических проблем является интеграция данных, наличие пропусков и ошибок и сложность стандартизации медицинских данных [42]. Если для омиксных данных ведущей проблемой является разнородность форматов и сложности интерпретации [33], то для данных из медицинских карт основные проблемы связаны с деятельностью медицинского персонала — пропуски в записях, ошибки ввода, некорректная интерпретация и др. Кроме этого, до сих пор не решены задачи стандартизации единиц измерения и норм для лабораторных и инструментальных данных и многие другие, которые затрудняют анализ даже не самых сложных массивов данных [42].

С точки зрения технологий анализа, в том числе технологий искусственного интеллекта, есть проблема применимости ряда алгоритмов машинного обучения к медицинским данным, а также большая проблема восприятия заключений, формируемых

на основе алгоритмов ИИ медицинскими работниками. Следом за этим встают юридические проблемы персональной ответственности за медицинские решения, принимаемые на основе таких алгоритмов, и иные вопросы права, касающиеся регулирования применения подобных систем в здравоохранении, которые подпадают под определение «изделие медицинского назначения».

Важной этической и организационной проблемой является защита персональных данных и безопасность данных [43–44]. В первом случае необходим четкий протокол доступа к персональным данным и мониторинг согласия пациента на предоставление к ним доступа, в том числе третьим лицам, что сегодня не всегда хорошо отработано в медицинских организациях. Во втором случае — глобальная защита данных от их скачивания и использования в нелегальных целях. Во всем мире и в России предпринимаются достаточно большие усилия по обеспечению безопасности, существует строгое законодательство, но, тем не менее, утечка данных и их использование, в том числе в корыстных или даже криминальных целях, остается реальным фактором риска.

В отношении анализа омиксных данных тоже существуют технические сложности, которые складываются из наличия «шума» в данных, необходимости стандартизации условий эксперимента, особенностей биообразцов, технологий и др. В связи с этим кажущаяся легкость объединения данных и применения алгоритмов анализа нередко сталкивается с необходимостью сложного процессинга данных, их «чистки» и устранения шумов, прежде чем они будут пригодны для реального анализа, визуализации закономерностей, кластеризации и классификации, что предполагает наука о больших данных. Еще одной сложностью омиксных данных является то, что число параметров (белков, генов, метаболитов), как правило, намного больше, чем число обследованных индивидуумов, что затрудняет построение классических связей и визуализацию данных, свойственных АБД [33].

Классические алгоритмы анализа больших данных включают [45]:

*Bayesian classification* — *Байесовский классификатор*;

*Neural network algorithm* — *нейросетевой алгоритм*;

*Decision tree induction* — *дерева решений*;

*Rule based classification* — *классификация, основанная на правилах*;

*Support vector machine* — *машина опорных векторов*;

*K-Nearest neighbor classifier* — *классификация по методу k-ближайших соседей*;

*Rough set approach* — подход, основанный на теории неточных множеств;

*Genetic algorithm* — генетический алгоритм;

*Fuzzy set approach* — подход, основанный на теории нечетких множеств.

Каждый алгоритм имеет свою область применения в машинном обучении, что определяется спецификой данных и поставленными задачами. Чаще всего используются первые три алгоритма, особенно в системах компьютерного зрения.

## ОБЛАСТИ УСПЕШНОГО ПРИМЕНЕНИЯ ТЕХНОЛОГИЙ БОЛЬШИХ ДАННЫХ И ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В БЛИЖАЙШЕМ БУДУЩЕМ

1. **Профилактика.** Считается, что анализ больших данных и соединение вместе омиксных данных и сведений по популяции и индивидуумам, с точки зрения поведения и факторов риска, приведет к существенному прогрессу в осознании механизмов развития заболеваний и формировании стратегий персонализированной профилактики. На этих данных также будет базироваться так называемая прецизионная профилактика, выделяющая группы риска на основании генотип-фенотипических взаимодействий.

2. **Экстренная медицина.** Анализ больших данных и ИИ могут уже в ближайшее время совершить революцию в неотложной терапии, позволяя на основании сбора данных с различного рода датчиков и устройств и интеграции данных с имеющейся информацией о пациенте, в том числе в социальных сетях и иных сферах, а также анализа изображений ускорить дифференциальный диагноз неотложных состояний и сформировать системы поддержки выбора тактики лечения.

3. **Управление популяционным здоровьем.** Прогнозирование динамики развития заболеваний, эпидемий, предиктивная и прескриптивная аналитика для планирования медицинской помощи и расходов.

4. **Мобильное здоровье (mobile health)** — мониторинг различных параметров здоровья и формирование поддержки клинических решений в пациент-центричной модели здравоохранения.

5. **Управление медицинской организацией,** регуляция потоков пациентов, предсказательное моделирование загрузки и работы персонала и оборудования. Повышение экономической эффективности медицинской помощи.

6. **Управление качеством** медицинской помощи и ее интегральная оценка. Снижение числа

ошибок и управление системой повышения квалификации медицинских работников.

7. **ИИ и анализ изображений.** Компьютерное зрение.

8. **Новые технологии drug-design.** Создание лекарств и проведение их испытаний на основе моделирования.

9. **Нанотехнологии и роботические технологии.**

10. **Формирование системы персонализированной медицины** на основе всех типов данных.

## ЗАКЛЮЧЕНИЕ

Анализ больших данных в медицине — это многообещающий инструмент, который позволит интегрировать между собой данные различных источников и систем, включая данные социальных сетей и носимых устройств. В результате будут развиваться новые отрасли знаний и компетенций, такие как биоинформатика, сенсорная информатика, вычислительная медицина и информатика здоровья. На этой основе будут развиваться новые сервисы и приложения, которые смогут управлять большими данными и строить на их основании клинические и организационные решения.

Персонализация медицины, сокращение сроков госпитализации и эффективность работы системы здравоохранения, в том числе снижение стоимости медицинской помощи, являются важнейшими целями модернизации здравоохранения, и они тесно связаны с цифровизацией и применением аналитики больших данных. Но внедрение новых технологий медицинской аналитики требует комплексного решения целого ряда проблем, начиная с технологических и заканчивая идеологией развития медицины в сторону уменьшения госпитальной помощи, вопросами этики и изменений законодательства. Безусловно, со временем налаживание системы обмена данными и построение предиктивной аналитики не только на уровне одного учреждения или конкретного пациента, а на уровне системы здравоохранения региона, страны сделает систему более устойчивой и прозрачной и будет способствовать повышению качества и безопасности медицинской помощи. Для успешного внедрения технологии БД и ИИ необходимо вовремя создавать инфраструктуру и готовить специалистов, способных воспринять инновации и быстро встроить их в клиническую работу [46–47].

## Конфликт интересов / Conflict of interest

Авторы заявили об отсутствии потенциального конфликта интересов. / The authors declare no conflict of interest.

## СПИСОК ЛИТЕРАТУРЫ / REFERENCES

1. Aceto G, Persico V, Pescapé A. The role of Information and Communication Technologies in healthcare: Taxonomies, perspectives, and challenges. *J. Netw. Comput. Appl.* 2018;107:125–154. DOI: 10.1016/j.jnca.2018.02.008.
2. Workgroup for Electronic Data Interchange. Maximizing the potential of genomic information to improve care coordination and health outcomes <https://www.wedi.org/docs/publications/a-white-paper-by-the-genomics-workgroup.pdf?sfvrsn=0>.
3. Fatt QK; Ramadas A. The Usefulness and Challenges of Big Data in Healthcare. *J. Healthc. Commun.* 2018, 3, 1–4.
4. Aarathi S; Vasundra S. Impact of healthcare predictions with big data analytics and cognitive computing techniques. *Int. J. Recent Technol. Eng.* 2019, 8, 4757–4762.
5. Lhotska L. Application of industry 4.0 concept to health care. *Stud. Health Technol. Inform.* 2020, 273, 23–37.
6. EMC digital universe with research & analysis by IDC Healthcare. <http://www.emc.com/analyst-report/digital-universe-healthcare-vertical-report-ar.pdf>).
7. IDC. Driving data growth in healthcare. <https://image.slidesharecdn.com/emckeyssuccessfactorsr ioitahealth2015-150829140335-lva1-app6892/95/healthcare-it-transformation-key-success-factors-2-638.jpg?cb=1440858718> )
8. Hansen MM, Miron-Shatz T, Lau AYS, Paton C. Big Data in Science and Healthcare: A Review of Recent Literature and Perspectives. Contribution of the IMIA Social Media Working Group. *Yearb. Med. Inform.* 2014, 9, 21–26.
9. Thuemmler C, Bai C. Health 4.0: How Virtualization and Big Data Are Revolutionizing Healthcare; Springer International Publishing; Cham, Switzerland, 2017; ISBN 9783319476179.
10. Lee I. Big data: Dimensions, evolution, impacts, and challenges. *Bus. Horiz.* 2017, 60, 293–303.
11. Chinnaswamy A, Papa A, Dezi L, Mattiacci A. Big data visualisation, geographic information systems and decision making in healthcare management. *Manag. Decis.* 2019, 57, 1937–1959.
12. Sumarsono, Anshari M, Almunawar MN. Big Data in Healthcare for Personalization Customization of Healthcare Services. In Proceedings of the 2019 International Conference on Information Management and Technology Jakarta/Bali, Indonesia, 19–20 August 2019; Volume 1, pp. 73–77.
13. Chen HC, Chiang RH. Business intelligence and analytics: From big data to big impact. *MIS Q.* 2012, 36, 1165–1188.
14. Tran TQB, du Toit C, Padmanabhan S. Artificial intelligence in healthcare—the road to precision medicine. *J. Hosp. Manag. Health Policy* 2021, 5, 29.
15. Weaver CA, Ball MJ, Kim GR, Kiel JM. Healthcare information management systems: Cases, strategies, and solutions: Fourth edition. In *Healthcare Information Management Systems: Cases, Strategies, and Solutions*, 4th ed.; Springer International Publishing: Cham, Switzerland, 2016; pp. 1–600. ISBN 9783319207650.
16. Firouzi F, Rahmani AM, Mankodiya K, Badaroglu M, Merrett GV, Wong P, Farahani B. Internet-of-Things and big data for smarter healthcare: From device to architecture, applications and analytics. *Future Gener. Comput. Syst.* 2018, 78, 583–586.
17. Agrawal D, Madaan J. A structural equation model for big data adoption in the healthcare supply chain. *Int. J. Product. Perform. Manag.* 2021.
18. Macpherson A, Holt R. Knowledge, learning and small firm growth: A systematic review of the evidence. *Res. Policy* 2007, 36, 172–192.
19. Choi TM. Incorporating social media observations and bounded rationality into fashion quick response supply chains in the big data era. *Transp. Res. Part E Logist. Transp. Rev.* 2018, 114, 386–397.
20. Kharbouch A, Naitmalek Y, Elkhokhi H, Bakhouya M, De Florio V, El Ouadghiri MD, Latre S, Blondia C. IoT and big data technologies for monitoring and processing real-time healthcare data. *Int. J. Distrib. Syst. Technol.* 2019, 10, 17–30.
21. Auffray C, Balling R, Barroso I, Bencze L, Benson M, Bergeron J, Bernal-Delgado E, Blomberg N, Bock C, Conesa A. Making sense of big data in health research: Towards an EU action plan. *Genome Med.* 2016, 8, 1–13.
22. Normandeau K. Beyond Volume, Variety and Velocity is the Issue of Big Data Veracity. *Inside Big Data* 2013. Available online: <https://insidebigdata.com/2013/09/12/beyond-volume-variety-velocity-issue-big-data-veracity/> (accessed on 18 January 2022).
23. Gandomi A, Haider M. Beyond the hype: Big data concepts, methods, and analytics. *Int. J. Inf. Manag.* 2015, 35, 137–144.
24. Diebold FX, Cheng X, Diebold S, Foster D, Halperin M, Lohr S, Mashey J, Nickolas T, Pai M, Pospiech M. A Personal Perspective on the Origin (s) and Development of “Big Data”: The Phenomenon, the Term, and the Discipline\*. 25. Dash S, Shakyawar SK, Sharma M, Kaushik, S. Big data in healthcare: Management, analysis and future prospects. *J. Big Data* 2019, 6, 54.
26. Wang Y, Hajli N. Exploring the path to big data analytics success in healthcare. *J. Bus. Res.* 2017, 70, 287–299.
27. Yin Y, Zeng Y, Chen X, Fan Y. The internet of things in healthcare: An overview. *J. Ind. Inf. Integr.* 2016, 1, 3–13.

28. Oliver N, Arnesh T, Tak I. Smart hospital services: Health 4.0 and opportunity for developing economies. In Proceedings of the Towards the Digital World and Industry X.0—Proceedings of the 29th International Conference of the International Association for Management of Technology, IAMOT 2020, Cairo, Egypt, 13–17 September 2020; pp. 345–361.
29. Wang Y, Kung LA, Byrd TA. Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. *Technol. Forecast. Soc. Chang.* 2018, 126, 3–13.
30. Nalluri S, Sasikala R. An insight into application of big data analytics in healthcare. *Int. J. Data Min. Model. Manag.* 2020, 12, 87–117.
31. Singh K, Jahnke I, Mosa A, Calyam P. The Winding Road of Requesting Healthcare Data for Analytics Purposes: Using the One-Interview Mental Model Method for Improving Services of Health Data Governance and Big Data Request Processes. *J. Bus. Anal.* 2021, 1–18.
32. Singh RK, Agrawal S, Sahu A, Kazancoglu Y. Strategic issues of big data analytics applications for managing health-care sector: A systematic literature review and future research agenda. *TQM J.* 2021.
33. Hassan M, Awan FM, Naz A, et al. Innovations in Genomics and Big Data Analytics for Personalized Medicine and Health Care: A Review. *Int. J. Mol. Sci.* 2022, 23, 4645. <https://doi.org/10.3390/ijms23094645>
34. Cirillo D, Valencia A. Big data analytics for personalized medicine. *Curr. Opin. Biotechnol.* 2019, 58, 161–167.
35. Lee J. *Industrial AI: Applications with Sustainable Performance*; Springer: Berlin/Heidelberg, Germany, 2020.
36. Martínez-Rojas M, del Carmen Pardo-Ferreira M, Rubio-Romero JC. Twitter as a tool for the management and analysis of emergency situations: A systematic literature review. *Int. J. Inf. Manag.* 2018, 43, 196–208.
37. Bates DW, Saria S, Ohno-Machado L, Shah A, Escobar G. Big data in health care: Using analytics to identify and manage high-risk and high-cost patients. *Health Aff.* 2014, 33, 1123–1131.
38. Zolbanin HM, Delen D, Sharma SK. The strategic value of big data analytics in health care policy-making. *Int. J. E-Bus. Res.* 2018, 14, 20–33.
39. Feldman B, Martin E, Skotnes T. Big data in healthcare: Hype and hope. *Dr. Bonnie 360 degree (Business Development for Digital Health)*. 2012. Available online: <http://www.riss.kr/link?id=A99883549> (accessed on 12 February 2022).
40. Basile LJ, Carbonara N, Pellegrino R, Panniello U. Business intelligence in the healthcare industry: The utilization of a data-driven approach to support clinical decision making. *Technovation* 2022, 102482.
41. Patel S, Patel A. A Big Data Revolution in Health Care Sector: Opportunities, Challenges and Technological Advancements. *Int. J. Inf. Sci. Tech.* 2016, 6, 155–162.
42. Murdoch TB, Detsky AS. The inevitable application of big data to health care. *JAMA-J. Am. Med. Assoc.* 2013, 309, 1351–1352.
43. Rumsfeld JS, Joynt KE, Maddox TM. Big data analytics to improve cardiovascular care: Promise and challenges. *Nat. Rev. Cardiol.* 2016, 13, 350–359.
44. Senthilkumar SA, Rai BK, Meshram AA, Gunasekaran A, Chandrakumarmangalam S. Big Data in Healthcare Management: A Review of Literature. *Am. J. Theor. Appl. Bus.* 2018, 4, 57–69.
45. Weng C, Kahn MG. Clinical Research Informatics for Big Data and Precision Medicine. *Yearb. Med. Inform.* 2016, 25, 211–218.
46. Tariq MA, Hoyle DC. Translating the Machine: Skills that Human Clinicians Must Develop in the Era of Artificial Intelligence. *Ophthalmol. Ther.* 2022, 11, 69–80.
47. De Mauro A, Greco M, Grimaldi M, Ritala P. Human resources for Big Data professions: A systematic classification of job roles and required skill sets. *Inf. Process. Manag.* 2018, 54, 807–817.

#### Информация об авторах:

Шляхто Евгений Владимирович, д.м.н., профессор, академик РАН, генеральный директор ФГБУ «НМИЦ им. В. А. Алмазова» Минздрава России;

Конради Александра Олеговна, д.м.н., профессор, академик РАН, заведующий кафедрой организации, управления и экономики здравоохранения Института медицинского образования, заместитель генерального директора по научной работе ФГБУ «НМИЦ им. В. А. Алмазова» Минздрава России;

Курापеев Дмитрий Ильич, заместитель генерального директора по информационным технологиям и проектному управлению ФГБУ «НМИЦ им. В. А. Алмазова» Минздрава России.

#### Author information:

Shlyakhto Evgeny V., MD, PhD, Professor, Academician of RAS, Director General of the Almazov National Medical Research Centre;

Konradi Aleksandra O., Doctor of Medical Sciences, Professor, Academician of the Russian Academy of Sciences, Head of the Department of Organization, Management and Economics of Healthcare of the Institute of Medical Education, Deputy Director General on Research of the Almazov National Medical Research Center;

Kurapeev Dmitry I., Deputy Director General for Information Technology and Project Management of the Almazov National Medical Research Centre.