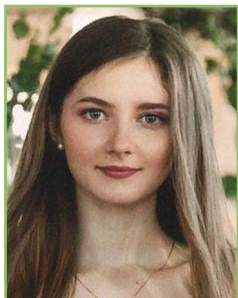


ОРИГИНАЛЬНАЯ СТАТЬЯ

УДК 004.6(045)  
© Иванова А. А., 2020

# Применение Big Data в сфере здравоохранения: российский и зарубежный опыт



**Алина Алексеевна Иванова**, студентка факультета экономики и финансов топливно-энергетического комплекса, Финансовый университет, Москва, Россия

**Alina A. Ivanova**, student of the Faculty of Economics and Finance of the Fuel and Energy Complex, Financial University, Moscow, Russia  
lina-twins@yandex.ru

## АННОТАЦИЯ

В статье проведен анализ применения больших данных в сфере здравоохранения. Описаны разные подходы к определению понятия «большие данные» (Big Data), сферы применения данной технологии; показаны наиболее востребованные технологии по отраслям. Рассмотрены предпосылки применения больших данных в здравоохранении, а также представлены достоинства и недостатки применения больших данных в здравоохранении. Отмечается исторический опыт применения Big Data в России и США, анализируются конкретные примеры применения в России, Европе, США, Китае. В статье проведен прогноз рынка больших данных в российском здравоохранении и показан практический пример использования Python в рамках анализа большого потока медицинских данных. Сделаны выводы о характерных тенденциях использования больших данных в различных странах. Делается вывод о важности применения больших данных в здравоохранении.

**Ключевые слова:** большие данные; здравоохранение; аналитика; патент; прогноз; проекты

**Для цитирования:** Иванова А. А. Применение Big Data в сфере здравоохранения: российский и зарубежный опыт. *Научные записки молодых исследователей*. 2020;8(5):42-53.

ORIGINAL PAPER

# Application of Big Data in Healthcare Entities: Russian and Foreign Experience

## ABSTRACT

The article analyses the use of big data in healthcare entities. Described are different approaches to defining the concept of Big Data, the scope of this technology; shows the most popular technologies by industry. The author considered the prerequisites for the use of big data in healthcare and presented the advantages and

Научный руководитель: **Долганова О.И.**, кандидат экономических наук, доцент кафедры «Бизнес-информатика», Финансовый университет, Москва, Россия / Scientific supervisor: **Dolganova O.I.**, Cand. Sci. (Economics), Associate Professor, Department of Business Informatics, Financial University, Moscow, Russia.

*disadvantages of using big data in healthcare. Also, the author noted the historical experience of using Big Data in Russia and the USA and analysed specific examples of application in Russia, Europe, USA, China. The article presents a forecast of the big data market in Russian healthcare. It shows a practical example of using Python in the analysis of a massive flow of medical data. The author has concluded the typical trends in the use of big data in different countries and about the importance of using big data in healthcare.*  
**Keywords:** Big Data; healthcare; analytic; patent; forecast; projects

**For citation:** Ivanova A. A. Application of Big Data in healthcare entities: Russian and foreign experience. *Nauchnye zapiski molodykh issledovatelei = Scientific notes of young researchers.* 2020;8(5):42-53.

## Введение

На сегодняшний день технологии больших данных проникли в разные отрасли, наиболее актуальной из которых является здравоохранение. Аналитики прогнозируют стремительный рост рынка больших данных, а целесообразность их применения в системе здравоохранения широко обсуждается мировым сообществом.

Целью исследования является рассмотрение рынка больших данных в здравоохранении. В раскрытие цели можно выделить следующие задачи:

1. Анализ теоретических основ Big Data.
2. Рассмотрение предпосылок, преимуществ и недостатков Big Data в здравоохранении.
3. Анализ российского и зарубежного опыта применения технологий больших данных в здравоохранении.
4. Демонстрация конкретного примера применения больших данных в здравоохранении и составление прогноза объема рынка.

В качестве основных методов исследования использовались методы сравнения, описания, аналитического выравнивания по прямой – для построения прогнозов.

## Big Data: теоретические основы и использование в здравоохранении

В связи с ускорением научно-технического прогресса и развитием нового поколения технологий появилась возможность работать с огромными объемами данных. Только в 2018 г. человечество накопило 33 зеттабайта данных<sup>1</sup>. Можно отметить, что 1 зеттабайт (равен  $10^{12}$  Гб) эквивалентен объему памяти одного человеческого мозга. Согласно определению Gartner, под «большими данными»

(Big Data) следует понимать большой объем, высокую скорость и/или большое разнообразие информационных активов, которые требуют экономически эффективных, инновационных форм обработки информации, позволяющих улучшить понимание, принятие решений и автоматизацию процессов<sup>2</sup>.

Впервые термин «Big Data» был использован в 2008 г. американским редактором журнала «Nature» Клиффордом Клинчем<sup>3</sup>: «большими данными» он назвал взрывной рост потоков информации. К таким потокам он отнес все массивы свыше 150 Гб в сутки.

Сравнительное преимущество аналитики больших данных по отношению к традиционной аналитике заключается в скорости обработки большого объема данных, отсутствии необходимости сортировки информации, возможности анализа поступающих данных в режиме реального времени. Наоборот, традиционная аналитика требует постепенной обработки, предварительного сортирования и редактирования [1, с. 256].

Свойства больших данных характеризуются как 3V: volume (объем), variety (разнообразие), velocity (скорость). Сферы использования больших данных разнообразны<sup>4</sup>:

1. Продуктовая разработка: используя данные из социальных сетей и результаты пробных продаж, принимается решение о выводе на рынок нового продукта. Например, компания Netflix

<sup>2</sup> Gartner Glossary. URL: <https://www.gartner.com/en/information-technology/glossary/big-data> (дата обращения: 12.04.2020).

<sup>3</sup> Словарь маркетолога. URL: <https://www.calltouch.ru/glossary/big-data/> (дата обращения: 12.04.2020).

<sup>4</sup> Большие данные (Big Data) в России. URL: [http://www.tadviser.ru/index.php/Статья:Большие\\_данные\\_\(Big\\_Data\)\\_в\\_России](http://www.tadviser.ru/index.php/Статья:Большие_данные_(Big_Data)_в_России) (дата обращения: 15.03.2020).

<sup>1</sup> URL: <https://news.rambler.ru/internet/41894466-mirovoy-obem-dannyh-prevysit-17-zettabayt> (дата обращения: 15.03.2020).

Таблица 1

Самые востребованные технологии по отраслям в 2019 г.

Технология	Суммарный рейтинг	Финансовая сфера	Торговля	Медицина	Нефтегазовый комплекс
Аналитика больших данных	83	91	85	65	73
Искусственный интеллект	72	75	58	76	56
Облачные решения	66	46	59	44	55

Источник: составлено автором на основе данных CNews Analytics, 2019. URL: <https://www.cnews.ru/reviews/ittrendy2020> (дата обращения: 15.03.2020).

создала сериал «Карточный домик» на основе анализа больших данных.

2. Прогнозирование: анализ больших данных позволяет выявить факторы, которые помогут предсказать сбои работы оборудования.

3. Машинное обучение: большие данные сделали возможным обучение машины, а не ее программирование.

4. Соблюдение нормативных требований: с помощью аналитики больших данных могут быстро выявляться мошеннические действия. Подобные решения активно внедряются в банках и налоговых органах.

Одной из наиболее социально значимых отраслей для применения Big Data является здравоохранение [1, с. 255]. Системы, позволяющие работать с Big Data, являются уникальным инструментом в решении задач здравоохранения, поскольку позволяют проанализировать большой поток информации о пациентах, особенностях возникновения и протекания болезней, действии препаратов; составлять прогнозы возникновения эпидемий; определять более эффективные методы лечения и т.д. Кроме того, использование инструментов работы с большими данными особенно актуально в здравоохранении ввиду роста запросов пациентов к качеству обслуживания [2, с. 46].

Направления использования Big Data в здравоохранении сводятся к возможности оказания персонализированной медицинской помощи, повышения уровня диагностики, профилактики эпидемий, эффективности лечения и борьбе с мошенниками в страховой медицине.

В табл. 1 представлены данные о том, какие современные технологии в каких отраслях востребованы. Опрос проводился в онлайн-режиме посредством его размещения на сайте CNews.ru. В опросе приняли участие представители отрасли здравоохранения и компании – заказчики медицинских услуг. Рейтинг каждой технологии рассчитывался как процент голосов респондентов, которые указали ее как одну из наиболее востребованных.

Из таблицы видно, что 65% опрошенных выбрали аналитику больших данных в здравоохранении самой востребованной технологией.

К большим данным в сфере здравоохранения можно отнести [2, с. 46]:

- данные, полученные из диагностики;
- данные с датчиков мониторинга;
- информацию, полученную при неотложной помощи;
- результаты применения нетрадиционной медицины;
- нормативные документы в области здравоохранения;
- данные о лекарствах и т.д.

Примером больших данных в здравоохранении можно считать всю информацию о генетических особенностях организма, что составляет сотни Гб на человека.

### Предпосылки развития Big Data в здравоохранении

Компания PricewaterhouseCoopers в отчете «Global top health industry issues: Defining the

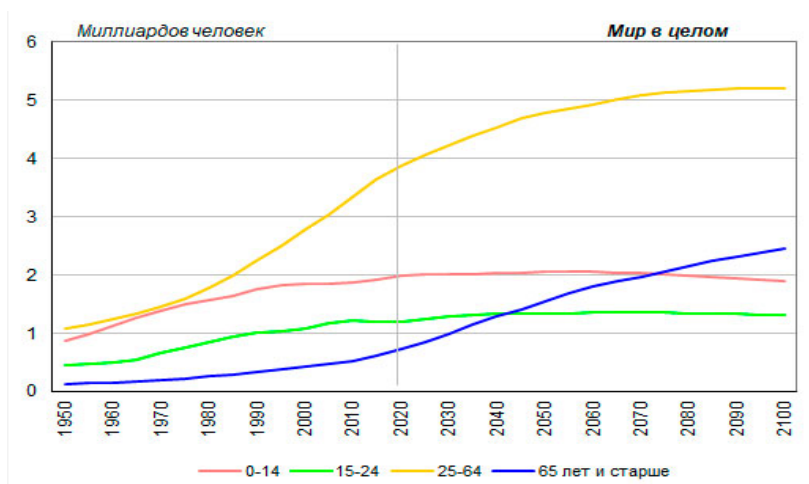


Рис. 1. Прогноз численности населения по возрастам

Источник: United Nations, Department of Economic and Social Affairs, Population Division 2019. URL: <http://www.demoscope.ru/weekly/2019/0837/barom01.php> (дата обращения: 15.03.2020).

healthcare of the future»<sup>5</sup> выделила 8 факторов, способствующих упрощению системы здравоохранения и ее активному прогрессу:

1. Использование искусственного интеллекта.
2. Составление четких инструкций для применения искусственного интеллекта.
3. Комфортные условия для пациента.
4. Преобразование клинических испытаний.
5. Использование технологий для распространения дистанционных услуг.
6. Формирование ценных медицинских технологий.
7. Кибербезопасность.
8. Социальные факторы – уровень дохода и образования и т.д.

Также, согласно отчету, можно выделить следующие глобальные тенденции в развитии здравоохранения: снижение доверия к институтам и технологиям (пациенты требуют больше информации о выборе лечения), старение населения. Значительная часть населения стареет и предъявляет все больше требований к местной инфраструктуре здравоохранения и к социальным учреждениям. Согласно данным ООН, в середине 2019 г. численность людей в возрасте 65 лет и старше составляла 703 млн человек, т.е. около 10% всего населения мира.

По прогнозу ООН численность населения в возрасте 65 лет и старше достигнет порядка 1,5 млрд человек к 2050 г. (рис. 1).

<sup>5</sup> Отчет компании PriceWaterhouse Coopers Global top health industry issues. 2018. URL: <https://www.pwc.com/gx/en/healthcare/pdf/global-top-health-industry-issues-2018-pwc.pdf> (дата обращения: 15.03.2020).

Перечисленные выше тенденции выступают факторами развития Big Data в здравоохранении, что, в свою очередь, призвано способствовать разрешению указанных проблем.

Кроме того, на настоящий момент наблюдается рост пользовательской активности в интернете, что способно стать полноценной предпосылкой активного развития Big Data в здравоохранении [3, с. 27]. Психология современного человека построена по принципу: «не иду к врачу, жалуюсь в социальных сетях». Потенциальный больной сначала спросит совета в интернете, только потом, возможно, пойдет к врачу. С помощью технологий больших данных можно было бы проанализировать тексты подобных сообщений и выявить возможные вспышки вируса.

Стоит отметить возникшую необходимость хранения данных о пациенте в одном месте. Применение анализа больших данных в данном случае позволит решить проблему неструктурированности медицинских данных, более точно определять методы лечения, побочные эффекты препаратов и т.д.

### Преимущества и недостатки применения Big Data в здравоохранении

Технологии Big Data способствуют упрощению некоторых процессов в здравоохранении. Например, с помощью анализа Big Data улучшится качество клинических испытаний. Кроме того, в качестве основных преимуществ применения таких технологий можно выделить [4, с. 9]:

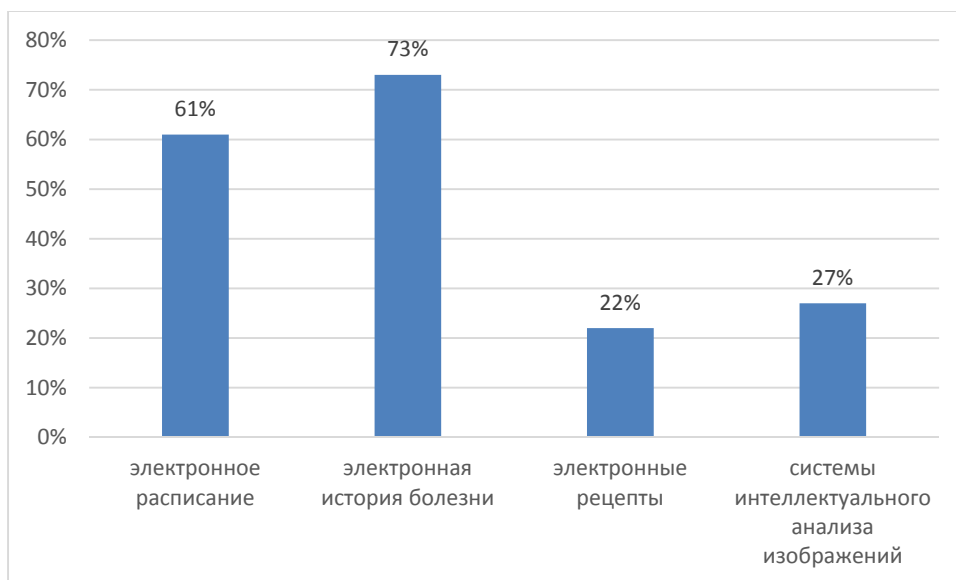


Рис. 2. Цифровые технологии, используемые в медицинской практике

Источник: URL: <https://ict.moscow/event/digital-pharma-2019/> (дата обращения: 15.03.2019).

1. Возможность самостоятельно контролировать свое здоровье.

2. Упрощение процесса принятия решения о диагнозе пациента ввиду компьютерного анализа анамнеза.

3. Переход от традиционных медицинских методов к совершенствующимся методам с сохранением накопленного опыта.

Стоит отметить, что огромное влияние на перспективы больших данных оказывает отношение самих врачей к изменениям.

В рамках российского интернет-проекта «Доктор на работе» проводится ежегодное исследование поведения врачей. В частности, в 2019 г. был проведен опрос врачей на тему перспектив Big Data в здравоохранении. В опросе приняли участие 1112 врачей, из которых 70% – сотрудники бюджетных учреждений.

По результатам опроса 93% врачей видят преимущества Big Data перед традиционными инструментами. На рис. 2 представлены проценты проголосовавших за часто используемые цифровые технологии. Из рисунка видно, что чаще всего используется электронная история болезни и электронное расписание.

В качестве основных преимуществ систем интеллектуального анализа изображений можно выделить: экономия времени (86% голосов), более точная диагностика (38%), более эффективное лечение (10%), снижение процента серьезных заболеваний (4%).

При этом лишь 53% врачей согласны на внедрение IT-системы с целью анализа данных историй болезни пациентов. Это связано с недостатками технологий сбора, хранения и обработки больших данных. IT-системы могут ошибаться и ломаться, а любая ошибка в здравоохранении стоит дорого, а также работа с новыми технологиями может занимать много времени.

Стоит отметить, что технологии Big Data требуют приватности, поэтому необходимо сохранение конфиденциальности. Доступ к некоторой информации даже может нанести ущерб. Примером может служить продажа информации о болезни для увеличения страховых взносов в связи с риском инфаркта<sup>6</sup>.

Внедрение программного обеспечения Big Data будет стоить дорого в здравоохранении, так как любая ошибка в программе или логике ее работы может стоить человеку жизни<sup>7</sup>.

### Зарубежный опыт использования Big Data в здравоохранении. Патенты

Зарубежные компании являются участниками российского рынка Big Data.

Компания Thomson Innovation проанализировала патентный рынок Big Data в здравоохранении (табл. 2).

<sup>6</sup> Кибербезопасность в здравоохранении: где недуг, где болезнь роста. URL: <https://www.kaspersky.ru/blog/healthcare-safeguarding-data/4474/> (дата обращения 15.03.2020).

<sup>7</sup> URL: <https://www.osp.ru/medit/2014/04/13040834.html> (дата обращения 11.04.2020)

Таблица 2

**Правообладатели российских патентов**

Патентообладатель	Количество документов
Koninkleike Philips Electronics NV	8
Dow Agrosiences LLC	1
Microsoft Corp	1
Ebay Inc	1
Mikhajlov Oleg Rostislavovich	1
Dialog Devices Ltd	1
Hoffmanco Internat OY	1

Источник: составлено автором на основе [4, с. 70].

По состоянию на 2016 г., из 14 выданных патентов только 1 принадлежит резиденту РФ – индивидуальному заявителю. Остальные патенты были получены зарубежными компаниями. За аналогичный период за рубежом было выдано 1810 патентов. Это говорит о том, что в России компании в 2010–2016 гг. еще не активно участвовали в проектах применения больших данных в здравоохранении [7, с. 71].

## Историческое развитие Big Data

### США

Мировым лидером по внедрению и применению технологий больших данных в здравоохранении является США [7, с. 70].

Актуальность применения больших данных в стране основывается на экономической эффективности. Ряд аналитиков уверены, что большие данные в здравоохранении создадут дополнительный финансовый поток 300 млрд долл. США преимущественно за счет снижения затрат на здравоохранение. Некоторые американские ученые полагают, что даже небольшой приток инвестиций в большие данные способен существенно повысить уровень жизни людей.

Еще в 90-е гг. в США отдельные медицинские общества стали заниматься созданием реестров и управлением ими. В 1997 г. Американский колледж кардиологов создал Национальный реестр сердечно-сосудистых данных (NCDR) для того, чтобы формализовать сбор данных и отчетность по диагностической катетеризации. С тех пор NCDR выросла до 10 различных реестров, которые содержат более 15 млн уникальных записей пациентов, охватывающих весь спектр сердечно-сосудистой помощи.

Недавно американский институт кардиологии (ACC) расширил сферу своих реестров до амбулаторного режима, создав два уникальных реестра: PINNACLE и Diabetes Collaborative Registry. PINNACLE – это крупнейший амбулаторный регистр повышения качества в кардиологии, в котором отслеживаются данные по поводу ишемической болезни сердца, гипертонии, сердечной недостаточности, а также 15 млн записей о пациентах. Данные из реестра PINNACLE автоматически передаются в систему отчетов о качестве врача. Целью недавно созданного Diabetes Collaborative Registry является объединение врачей первичной и специальной помощи с целью улучшения ухода

за больными и лечения сахарного диабета. Непрерывное расширение и использование данных реестра приведет к существенным изменениям в оказании медицинской помощи.

Калифорнийские ученые доказали, что элементарный анализ данных из социальных сетей способен выявить поведение, провоцирующее появление ВИЧ, создать систему мероприятий для борьбы с эпидемиями.

Тот же калифорнийский университет внедрил технологии больших данных в терапию местных больниц. Система запрограммирована на постоянное отслеживание изменения состояния больного, рассчитывает риск возникновения сепсиса.

Бостонские больницы также применяют технологии больших данных. При проведении серьезных операций прогнозируется риск летального исхода. Иногда на основе предложенной системой аналитики принимается решение о переносе операции на другой день с целью стабилизации показателей пациента [5, с. 68].

В числе первых проектов в США – Oncospace, или умная база данных. В такой базе хранятся все снимки раковых опухолей шеи, головы, поджелудочной железы, информация об анамнезе самих пациентов. Результаты применения данного проекта показали, что повышается эффективность и безопасность программ лучевой терапии [1, с. 256].

В качестве отдельного направления применения больших данных стоит выделить исследование ДНК человека. В США существует компания Affymetrix, занимающаяся разработкой чипов.

Такой биочип способен анализировать сотни и тысячи клеточных процессов одновременно.

Развитие больших данных в американском здравоохранении обусловлено большим притоком инвестиций. Еще президент Обама подтвердил важность «больших данных» в геномике, когда он объявил в своем выступлении в 2015 г. о выделении 215 млн долл. США на развитие многоуровневых баз данных в целях стимулирования междисциплинарного сотрудничества с особым акцентом на геномику [8, с. 24]. В США ожидается, что крупные капитальные вложения со стороны государственных и частных организаций приведут к росту знаний и применению геномики.

Аналитика больших данных используется не только для мониторинга отдельных людей, но и для исследований. Исследование Health eHeart, проведенное Университетом Калифорнии и Сан-Франциско, направлено на выявление моделей прогнозирования сердечных заболеваний, выявление причин фибрилляции предсердий, уменьшение числа случаев госпитализации при сердечной недостаточности и определение воздействия социальных сетей на здоровье сердца путем анализа до 1 млн участников в течение 10 лет.

Исследование будет использовать аналитику больших данных, чтобы ответить на эти вопросы с помощью метрик в реальном времени, полученных с помощью датчиков, надетых на пациента, мобильных приложений, социальных сетей и специального веб-портала.

Таким образом, в США развитие больших данных в здравоохранении идет полным ходом благодаря экспоненциальному росту доступных данных. Успех аналитики больших данных зависит от устойчивых технологических достижений в области информационных технологий и компьютерной архитектуры, а также от беспрепятственного сотрудничества и открытого обмена данными между врачами, плательщиками страхования, промышленностью и правительством.

### Европа

Как известно из данных Всемирной организации здравоохранения, Франция находится на первом месте в мире по уровню развития здравоохранения<sup>8</sup>.

<sup>8</sup> URL: <https://tranio.ru/france/healthcare/> (дата обращения 12.04.2020).

В сфере применения Big Data во Франции в 2017 г. была разработана система данных SNDS, которая объединяет информацию о здоровье человека, полученную из различных источников. Данная система получила распространение не только во Франции, но и во всей Европе. Ее уникальность состоит в возможности объединения информации о смертельных случаях, информации об инвалидности, баз данных страховых компаний. Доступ к SNDS имеют государственные службы и институты в области здравоохранения. Безопасность данных в системе обеспечивается за счет французского интернет-регулятора (CNIL), действие которого опирается на законодательство об охране персональных данных<sup>9</sup>.

Согласно прогнозам, интернет вещей ждет большая перспектива создания стоимости в здравоохранении (быстрое выявление заболеваний, лечение хронических болезней и т.д.) — порядка 16–35 млрд евро в 2020–2025 гг. [9, с. 131].

В Норвегии также уже используются системы больших данных в здравоохранении. Университетский госпиталь в столице работает с применением системы Power BI, позволяющей быстрее собрать данные и повысить точность исследований. Все это способствует перераспределению средств на более приоритетные медицинские задачи [10].

### Китай

Китай отличается от остальных стран тем, что уже начал строить специальные центры по анализу больших данных. В частности, ведется строительство государственного центра больших данных в области здравоохранения. Объем инвестиций в создание такого центра составил 1 млрд долл. США.

На настоящий момент в Китае ведется строительство крупного центра по ДНК-секвенированию. Центр будет использован для сбора информации о генофонде этнических китайцев. Реализацией проекта занимаются государственная компания Yangzi Group, Юго-Восточный университет и Нанкинский медицинский университет. В центре также планируется изучение

<sup>9</sup> Юркина Е. Национальная система данных о здоровье создана во Франции. URL: <https://habr.com/ru/post/328836/> (дата обращения 12.04.2020).

генетических мутаций, характерных для различных заболеваний<sup>10</sup>.

Таким образом, внедрение больших данных в здравоохранении по-прежнему зависит от значительного сбора данных и своевременного анализа для определения наиболее подходящей основы для оптимизации здоровья, индивидуальной профилактики, диагностики и лечения заболеваний. Большинство проектов по применению больших данных в здравоохранении направлены на изучение генома человека и решение проблемы своевременного выявления онкологии.

Обобщая все вышесказанное, стоит отметить, что на настоящий момент реализацией проектов больших данных в российском здравоохранении в основном заняты университеты и научно-исследовательские центры. В США субъектами являются не только университеты, но и компании, оказывающие инвестиционную поддержку. В Европе помимо университетов, в подобных проектах заинтересованы и государственные службы. Китай, скорее всего, находится на первом месте по уровню развития подобных проектов – государственные корпорации совместно с университетами активно развивают большие данные в здравоохранении.

### Российский опыт внедрения Big Data в здравоохранении

Мысли о возможностях применения больших данных в российском здравоохранении впервые появились в 70-х гг. прошлого столетия. Советские ученые занимались разработкой математических методов медицинской диагностики и прогнозирования. Ключевую роль в развитии исследований сыграл созданный в 1974 г. научный совет по медицинской кибернетике и вычислительной технике при Министерстве здравоохранения РСФСР. Под руководством научного руководителя С.А. Гаспаряна сотрудники совета занимались разработкой программ информатизации здравоохранения. Разработки велись совместно с отделом статистики и информатики Министерства здравоохранения РСФСР. Исследования можно считать значимыми ввиду географического охвата научных учреждений, центров, институтов:

были привлечены научно-исследовательский институт имени Склифосовского, государственный медицинский институт имени Пирогова, первый медицинский институт имени Сеченова и т.д. [5, с. 67].

Одним из значимых результатов проделанной работы является система, созданная для диагностики меланомы (злокачественной опухоли кожи). Разработанная система позволяет отличить меланому от доброкачественных и предзлокачественных образований по данным осмотра.

*Российский рынок больших данных в здравоохранении развивается преимущественно за счет научно-исследовательских центров и институтов, а компаний, реализующих свои решения в реальной практике, довольно мало.*

В последние годы в России выделяются две группы, занимающиеся разработкой Big Data в здравоохранении. **Первая группа** входит в состав информационно-аналитического агентства Департамента здравоохранения Москвы – это Центр анализа данных. К реализованным проектам Центра можно отнести проект «Кадровая аналитика». В нем представлена информация о работающих врачах и различных показателях их деятельности [5, с. 68].

Кроме того, на лекции представителя Центра анализа данных Е. Прокофьевой 18 февраля 2020 г. было озвучено, что в качестве плановых проектов могут быть рассмотрены прогнозирование эпидемий в школах, балансировка терапевтических участков, распознавание рака легких и т.д. Все проекты реализуются в рамках концепции «Умного города». По мнению центра региональной науки Венского технологического университета, Умный город – город с условиями для роста человеческого капитала. Для Умного города характерна человекоцентричность, доступность и открытость, защита персональных данных и т.д. Анализ медицинских данных в рамках города рассматривается как ступень реализации концепции «Умный город».

<sup>10</sup> URL: [http://zdrav.expert/index.php/Компания: National\\_Health\\_&\\_Medicine\\_Big\\_Data\\_\(Nanjing\)\\_Center](http://zdrav.expert/index.php/Компания: National_Health_&_Medicine_Big_Data_(Nanjing)_Center) (дата обращения 12.04.2020).



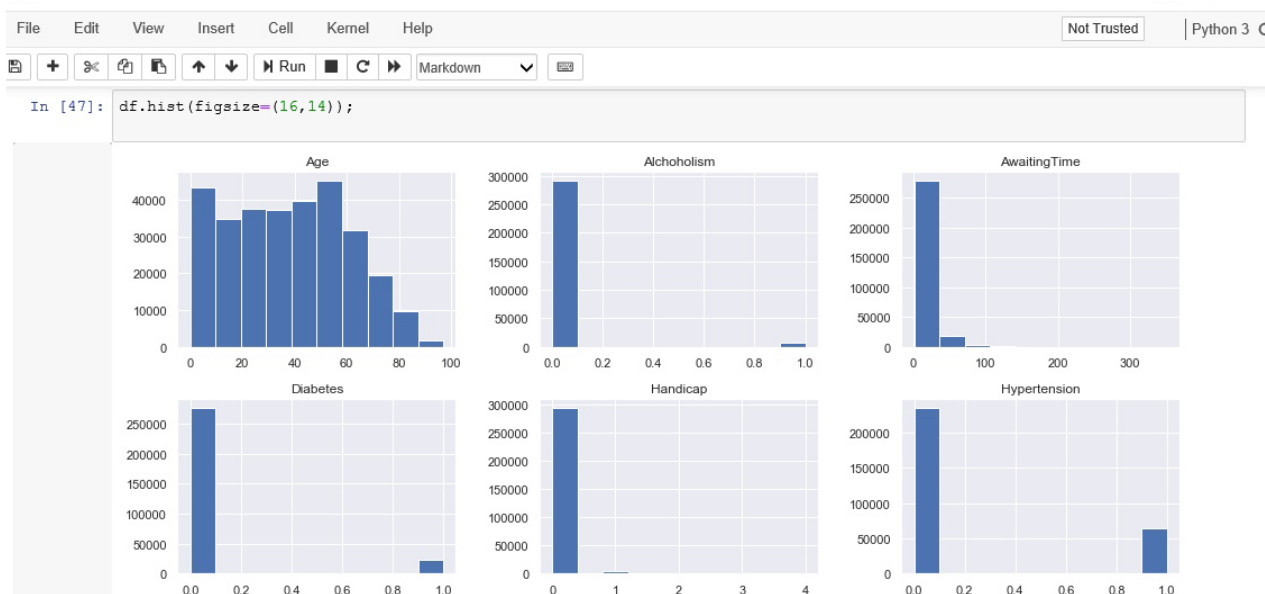


Рис. 3. Результат визуализации данных по пациентам Москвы

Источник: составлено автором с помощью Python и Anaconda Navigator на основе тестовой выборки, равной 300 000 чел.

**Вторая группа** входит в состав Департамента информационных технологий г. Москвы. Группа занимается аналогичными исследованиями на городском уровне.

*Важно расширять область применения больших данных, поскольку это позволит оптимизировать затраты времени и материальных ресурсов на разработку новых подходов к лечению, повысить качество подготовки медицинских работников. Использование больших данных – ключ к развитию превентивных мер в здравоохранении.*

В качестве российского опыта также можно рассмотреть дорожную карту «Хелснет»<sup>11</sup> Национальной технологической инициативы, инициированной Министерством здравоохранения Российской Федерации.

<sup>11</sup> План мероприятий («дорожная карта») «Хелснет» Национальной технологической инициативы. URL: [https://nti2035.ru/markets/docs/DK\\_healthnet.pdf](https://nti2035.ru/markets/docs/DK_healthnet.pdf) (дата обращения 15.03.2020).

В соответствии с дорожной картой «Хелснет» в 2017 г. был достигнут ряд контрольных показателей в области Big Data. В частности, была разработана программная платформа Health Neuristics с использованием технологий больших данных. Платформа используется для сбора биометрических данных спортсменов с целью организации тренировочных программ.

Российский рынок характеризуется большим разнообразием ИТ-компаний, специализирующихся на применении технологий больших данных. Такие компании имеют большой потенциал для выхода на рынок здравоохранения. Например, Яндекс, АБВУУ, Kaspersky – лидеры в области защиты персональных данных.

В соответствии с дорожной картой «Хелснет» в 2020–2025 гг. планируется развитие Big Data в здравоохранении, ожидается появление первых профессиональных центров, использующих программные продукты с целью непрерывного мониторинга состояния организма. В 2026–2035 гг. планируется создание уже целой сети таких центров.

Стоит отметить, что в России начинаются только первые шаги по использованию технологий больших данных в сфере здравоохранения.

В 2018 г. администрация Томской области совместно с российской компанией Rubius начали проект по анализу больших медицинских данных. В 2012 г. все медицинские учреждения Томской области подключились к единой медицинской

Таблица 3

## Результат визуализации данных по пациентам Москвы

	Age	Diabetes	Alcoholism	AwaitingTime
Mean	37,765	0,07	0,02	13,859

Источник: составлено автором с помощью Python и Anaconda Navigator на основе тестовой выборки, равной 300 000 чел.

Таблица 4

## Расчет объема рынка Big Data в российском здравоохранении

	Госрасходы, млрд руб.	Госрасходы на здравоохранение, млрд руб.	Доля, %	Объем рынка Big Data, млрд долл. США	Объем рынка Big Data в здравоохранении, млрд долл. США
2014	14 831,6	535,5	3,61	0,3	0,011
2015	15 620,3	516	3,30	0,5	0,017
2016	16 416,5	506,3	3,08	0,8	0,025
2017	16 420,3	439,8	2,68	1,1	0,029
2018	16 713	537,3	3,21	1,7	0,055
2019	18 293,7	678	3,71	2	0,074

Источник: составлено автором на основе данных [6].

информационной системе (МИС)<sup>12</sup>. Помимо МИС, к источникам данных также относятся база территориального фонда обязательного медицинского страхования и данные о вызовах электронной помощи. Проект оказался успешным, на настоящий момент планируется обучение специалистов социальной сферы работе с большими данными.

В феврале 2019 г. на ежегодном форуме «Big Data» был представлен ряд пилотных проектов в этой области. Например, система IBM Watson Explorer была обучена выявлять расхождения в протоколах описаний и заключений по пациентам, прошедшим лучевую диагностику. Точность такой системы составляет 95%.

Ярким примером применения Big Data в российском здравоохранении следует считать помощь в прикреплении к стоматологической поликлинике. Начиная с 1 января 2017 г. стоматологические поликлиники перешли на подушевое финансирование. Проблем с прикреплением людей, посещающих поликлинику по месту жительства

и регулярно, не было. Однако стояла проблема прикрепления людей, посещающих поликлинику не по месту жительства. Эту проблему решили технологии Big Data, с помощью них анализировалась место реального нахождения мобильного телефона [5, с. 67].

Современными инструментами анализа больших данных являются языки программирования такие, как R, Python. Используя Python в сочетании с Anaconda Navigator, можно проанализировать поток медицинских данных на предмет посещаемости поликлиник. Результат такого анализа потока медицинских данных 300 000 человек представлен на *рис. 3*.

Из загруженных данных по графикам можно сделать вывод, что чаще всех посещают поликлиники люди в возрасте 60–70 лет, при этом средний возраст посещения поликлиники составил 38 лет (*табл. 3*).

На 300 000 человек среднее число диабетиков составляет 0,07 человек, больным алкоголизмом – 0,02. Среднее время ожидания приема составляет 14 минут.

Таким образом, российский рынок больших данных в здравоохранении развивается преиму-

<sup>12</sup> URL: <https://www.riatomsk.ru/article/20180813/big-data-po-medicine-pomogut-spasti-tomichej-ot-infarktov-i-insuljotov/> (дата обращения 15.03.2020).

Расчетная таблица

Год	Объем рынка Big Data в здравоохранении, млрд долл. США	Теоретический уровень ряда динамики	Теоретический уровень ряда динамики в квадрате	Объем рынка Big Data в здравоохранении в соответствии с уровнем ряда динамики, млрд долл. США
2014	0,011	-5	25	-0,054
2015	0,017	-3	9	-0,049
2016	0,025	-1	1	-0,024
2017	0,029	1	1	0,029
2018	0,055	3	9	0,164
2019	0,074	5	25	0,371
Итого	0,210	0	70	0,436

Источник: составлено автором на основе данных [7, с. 70].

щественно за счет научно-исследовательских центров и институтов, а компаний, реализующих свои решения в реальной практике, довольно мало. По словам Вероники Скворцовой, бывшего министра здравоохранения России, официально заявленный план состоит в том, чтобы начать полноценное использование Big Data с 2020 г.<sup>13</sup>

### Прогнозы

По прогнозам компании Frost&Sullivan, рынок больших данных в здравоохранении станет одним из самых быстрорастущих<sup>14</sup>.

Точной информации об объемах этого рынка нет, но можно предположить, что он будет развиваться такими же темпами, как и весь рынок Big Data.

Рассмотрим примерные прогнозы объемов рынка Big Data в России. Для упрощения, путем умножения доли расходов на здравоохранение на объем рынка Big Data была получена доля здравоохранения в общем объеме рынка Big Data.

Методом аналитического выравнивания по прямой (табл. 4) можно построить прогноз роста объема рынка больших данных в России.

<sup>13</sup> URL: <https://tass.ru/obschestvo/4113127> (дата обращения 15.03.2020).

<sup>14</sup> URL: <http://www.tadviser.ru/index.php/> (дата обращения 15.03.2020).

$$\text{Уравнение тренда: } y = a_0 + a_1 \times t, \text{ где } a_0 = \frac{\sum y}{n} = \frac{0,210}{6} = 0,035, a_1 = \frac{\sum y \times t}{t^2} = \frac{0,436}{70} = 0,006.$$

Итоговое уравнение тренда:  $y = 0,035 + 0,006 \times t$ . Таким образом, прогноз на 2020 г.:  $0,035 + 0,006 \times 7 = 0,077$  млрд долл. США, на 2021 г. – 0,089 млрд долл. США.

В среднем рынок Big Data по объему характеризуется высокими темпами прироста – около 47%.

На примерах выше было продемонстрировано использование технологий больших данных в российском здравоохранении. Сегодня становится явной необходимость решать задачи, повышающие эффективность организации и осуществления лечебного процесса. В России направление больших данных только начинает развиваться, используя передовой зарубежный опыт.

Достижение эффективного управления медицинскими данными будет иметь важное значение для будущих систем здравоохранения [8, с. 25].

### Выводы

Вышеприведенные примеры показывают, что важно расширять область применения больших данных, поскольку это позволит оптими-

зировать затраты времени и материальных ресурсов на разработку новых подходов к лечению, повысить качество подготовки медицинских работников. Использование больших данных — ключ к развитию превентивных мер в здравоохранении.

### Список источников

1. Погосян И.А. Преимущества применения технологии Big Data в медицине. Образование и наука в современных реалиях. Чебоксары; 2019:255–257.
2. Малышева Ю.В. Big Data в здравоохранении. Вопросы современной науки: новые достижения. Самара; 2017:45–48.
3. Ефимова С.П., Нусина А.Ю. Большие данные в медицине. *NAUKA-RASTUDENT.RU*. 2016;(7):27–28.
4. Ижунин М.А. Big Data в здравоохранении. *Молодой ученый*. 2019;(50):8–10.
5. Ким С. Big Data в здравоохранении. *Московская медицина*. 2017;(1):65–69.
6. Christopher O. Austin. The application of Big Data in medicine: current implications and future decisions. *Journal of Interventional Cardiac Electrophysiology*. 2016(47):51–59.
7. Цветкова Л.А., Черненко О.В. Технология больших данных в медицине и здравоохранении России и мира. *Врач и информационные технологии*. 2016;(3):60–73.
8. Roberta Pastorino. Benefits and challenges of big data in healthcare: an overview of the European initiatives. *European Journal of Public Health*. 2019;(29):23–27.
9. Zhilina I. Yu. Big data et objets connectés. Faire de la France un champion de la révolution numérique. P.: Institut Montaigne. 2015;(1):126–132.
10. Рапаков Г.Г., Банщикова Г.Т. Интеллектуальный анализ данных в здравоохранении региона (на материалах Вологодской области). Монография. Вологда: ВоГУ; 2014.

### References

1. Poghosyan I.A. The benefits of using Big Data technology in medicine. *Obrazovaniye i nauka v sovremennykh realiyakh*. 2019:255–257. (In Russ.).
2. Malysheva Yu.V. Big Data in healthcare. *Voprosy sovremennoy nauki: novyye dostizheniya*. 2020: 45–48. (In Russ.).
3. Efimova S.P., Nusina A. Yu. Big data in medicine. *NAUKA-RASTUDENT.RU*. 2016;(7):27–28. (In Russ.).
4. Izhunin M.A. Big Data in healthcare. *Molodoy ucheny*. 2019;(50):8–10. (In Russ.).
5. Kim S. Big Data in Healthcare. *Moskovskaya meditsina*. 2017;(1):65–69. (In Russ.).
6. Christopher O. Austin. The application of Big Data in medicine: current implications and future decisions. *Journal of Interventional Cardiac Electrophysiology*. 2016(47):51–59.
7. Tsvetkova L.A., Chernenko O.V. Big data technology in medicine and healthcare of Russia and the world. *Vrach i informatsionnyye tekhnologii*. 2016;(3):60–73. (In Russ.).
8. Pastorino Roberta. Benefits and challenges of big data in healthcare: an overview of the European initiatives. *European Journal of Public Health*. 2019;(29):23–27.
9. Zhilina I. Yu. Big data and internet of things technologies: france prospects. Big data et objets connectés: Faire de la France un champion de la révolution numérique. P.: Institut Montaigne, 2015;(1):126–132. URL: <https://cyberleninka.ru/article/n/2016-01-038-tehnologii-bolshih-dannyh-i-interneta-veschey-perspektivy-frantsii-big-data-et-objets-connect-s-faire-de-la-france-un-champion-de-la>. (In Russ.).
10. Rapakov G.G., Bانشchikov G.T. Data mining in the region's healthcare system (based on the materials of the Vologda Oblast). Monograph. Vologda: VoGU; 2014.