

УДК 004.62

Мамедова М.Г.

Институт Информационных Технологий НАНА, Баку, Азербайджан
masuma.huseyn@iit.ab.az

BIG DATA В ЭЛЕКТРОННОЙ МЕДИЦИНЕ: ВОЗМОЖНОСТИ, ВЫЗОВЫ И ПЕРСПЕКТИВЫ

Рассмотрены факторы, определяющие взрывной рост информации в медицине. Исследованы специфические особенности медицинских данных, определены сущность феномена Big Data и его потенциал в электронной медицине. Проведена систематизация приложений Big Data и показаны возможности последних в поддержке принятия лечебно-диагностических и управленческих решений. В контексте развития и трансформации медицины выделены преимущества Big Data аналитики в этой сфере. Указаны вызовы, ограничивающие использование Big Data в медицине, и приведены перспективные направления их внедрения в исследуемую отрасль.

Ключевые слова: Big Data, электронная медицина, медицинские данные, Big Data приложения, принятие решений, Big Data вызовы.

Введение

С появлением и стремительным развитием информационно-коммуникационных и сетевых технологий мир столкнулся с лавинообразным нарастанием информации, получившим в обществе название «информационный взрыв» [1]. Количество создаваемой информации растет в геометрической прогрессии, причем примерно 90% данных в современном мире было создано за последние два года [2]. В соответствии с результатами нового исследования Digital Universe, проведенного аналитической компанией International Data Corporation (IDC) [3], за ближайшие 8 лет количество данных в мире достигнет 40 зеттабайт. Это значит, что на каждого жителя Земли будет приходиться по 5200 гигабайт данных. При этом, по оценкам IDC, на сегодняшний день проанализировано менее 1% всей имеющейся информации, а защищено – менее 20%.

Беспрецедентный рост информации, лишь ничтожная часть которой используется во всем мире, наглядно демонстрирует понимание того неоспоримого факта, что эффективное использование этого растущего количества информации должно стать одной из величайших научных и технических задач XXI столетия [4].

Медицина является одной из отраслей, исторически генерирующих большое количество данных, традиционно управляемых с помощью учета, соблюдения нормативно-правовых требований и критериев качества ухода за пациентами [5]. В то же время медицина относится к разряду сфер деятельности, в наибольшей степени отстающих от ряда клиенториентированных отраслей (банковское дело, ритейл и т.п.) в использовании ИТ и, соответственно, Big Data (больших данных). Впервые медицина обратилась к информационным технологиям в 60-х годах прошлого века для сбора и обработки больших объемов различных статистических данных. Однако количество медицинских данных значительно увеличилось лишь за последние 15–20 лет в связи с переходом отрасли на цифровой формат. Несмотря на хранение значительной части медицинских данных пока в бумажном виде, наблюдаемая тенденция к быстрой оцифровке уже способствовала накоплению терабайтов информации [4]. По некоторым расчетам объем медицинских данных в 2012 году достиг примерно 500 петабайт [6]. Последние исследования показывают, что более 30% всех данных, хранящихся на земле, составляют медицинские данные, и в будущем ожидается быстрое увеличение этой доли [7]. В соответствии с прогнозами к 2020 году количество медицинских данных будет достигать 25000 петабайт [8].

Огромные потоки медицинских данных предоставляют широкие возможности для развития методов и приложений по расширенному анализу последних. Естественно, реальная стоимость этого потока может быть понята только в случае содействия извлеченной из данных информации улучшению качества медицинских услуг [9, 10].

Целью настоящей статьи является исследование потенциала феномена Big Data в медицине как средства повышения качества медицинских услуг за счет возможностей аналитической поддержки принимаемых лечебно-диагностических и организационно-управленческих решений.

Ключевые элементы, определяющие информационный взрыв в медицине

Согласно [11] медицинские данные генерируются, в основном, за счет следующих трех ключевых элементов: совокупности персональных электронных медицинских записей (Electronic Health Records, EHR), биотехнологий, заложивших основу персонализированной медицины, и научных исследований и разработок (R&D). Этот список дополнен еще двумя бесспорными ключевыми элементами, обеспечивающими сегодня генерацию и стремительный рост информации, к которым относятся современные информационно-коммуникационные технологии (ИКТ) и Интернет.

EHR (электронная карта здоровья, ЭКЗ) [12] интегрирует в себе электронные персональные медицинские записи, относящиеся к одному человеку, собираемые и используемые несколькими медицинскими организациями. В русскоязычной литературе аналогом EHR является интегрированная электронная медицинская карта пациента (ИЭМК) [13]. В сущности ИЭМК пациента консолидирует накопленные в электронных медицинских картах (ЭМК) отдельных медицинских организаций персональные медицинские записи пациента. Термин ЭМК является аналогом международного термина Electronic Medical Record (EMR) и представляет собой электронную версию историй болезни пациентов.

Биотехнологии, вызвавшие бум в исследованиях молекулярной медицины, стимулировали бурный рост объема данных, особенно персональных геномных, в биологии и медицине. Дальнейшее развитие науки и технологий, обусловившее экспоненциально убывающие затраты на полногеномный анализ, способствовало взрывному росту данных, специфичных для определенного индивида. Так, если в 2003 году расшифровка одного человеческого генома стоила около 40 миллионов долларов, то в 2012 году стоимость определения последовательности генома человека упала ниже 1 тысячи долларов. В итоге эта процедура стала доступной широким массам [14]. Поскольку нарушение структуры генов является одной из причин возникновения какой-либо болезни, то геновая терапия может способствовать повышению эффективности назначенного лечения путем снижения негативных воздействий на организм. С помощью технологий Big Data можно сравнить генетический код человека с другими медицинскими показателями и выяснить, как именно его гены коррелируют с болезнями. Изучение генетических особенностей отдельного человека является одним из основополагающих элементов персонализированной медицины, главной целью которой является предоставление конкретному пациенту подходящего именно ему лекарственного препарата в правильных дозах и в конкретное время [15, 16].

К источникам Big Data относятся также другие биометрические данные: отсканированные отпечатки пальцев, почерка, сетчатки глаза, рентгеновские и другие медицинские изображения, значения жизненно важных показателей (артериального давления, пульса и т.д.).

(R&D). Медицина относится к одной из тех сфер, где наблюдается стремительное развитие науки и технологий, причем результаты научных открытий и разработок проявляются достаточно явно. В настоящее время мир сталкивается с новой проблемой –

экспоненциальным ростом скорости открытия медицинских знаний. На сегодня в каталоги биомедицинской литературы внесено более 18 миллионов статей, причем более 800 тысяч из них каталогизированы в 2008 году. Скорость пополнения медицинской литературы удваивается каждые 20 лет, и уже начиная с 2012 года ежегодное количество поступлений превышает 1 миллион [17].

Анализ динамики последних 150 лет, в течение которых эффекты индустриализации и информационного взрыва проявились наиболее остро, показал, что типичный врач общей помощи должен быть в курсе примерно 10 тысяч заболеваний и синдромов, 4 тысяч лекарственных препаратов, между которыми имеют место более 2 тысяч взаимодействий, определяющих возможность их совместного использования, 1,1 тысячи лабораторных анализов [18].

Технологии обработки больших данных могут оказать помощь в систематическом анализе неограниченных объемов разнотипной медицинской информации.

ИКТ и Интернет. Всё растущие темпы внедрения ИКТ в медицину не только дали мощный импульс к преобразованиям в этой сфере, но и становятся ключевыми инструментами в ее трансформировании. При этом наблюдается стремительное развитие самих ИКТ, сопровождаемое появлением новых платформ, аппаратных и программных продуктов, сетевых технологий, моделей сбора, хранения, обработки и анализа информации, с одной стороны, а также формируются все новые источники генерирования информации – с другой.

Несомненно, Интернет является важнейшим источником медицинской информации, предоставившим людям не доступную им ранее возможность приобретения знаний относительно болезней и медикаментов, уточнения диагноза, поиска как эффективного лечения, так и врача соответствующей специализации и т.п. Согласно данным проведенного в 2010 году Международным консалтинговым и маркетинговым агентством Manhattan research LLC исследования, половина опрошенных пациентов-респондентов использует Интернет для поиска медицинской информации о здоровье, заболеваниях и их профилактике, лекарственных средствах и т.п. [19]. При этом 9 из 10 пациентов принимают решение о собственном лечении, основываясь в том числе и на данных из сети. В то же время Интернет играет роль базовой инфраструктуры в процессах генерирования, быстрой, надежной и безопасной передачи медицинской информации.

По данным компании Cisco, в 2012 году из 2,8 зеттабайта передаваемой информации треть пришлось на долю данных, автоматически сгенерированных оборудованием и устройствами, присоединенными к Интернету или к Интернету вещей. По прогнозам Cisco IBSG, к 2020 году к Интернету будет подключено 50 миллиардов устройств, при том что эти прогнозы не учитывают ускоренного развития интернет-технологий и устройств [20]. С развитием Интернета вещей (*Internet of Thing*) и его дальнейшей эволюцией в Интернет всего (*Internet of Everything*), позволяющих связать между собой машины (*machine-to-machine, M2M*), появился еще один мощный источник генерирования данных [21]. Это показания объединенных в сети носимых устройств (дистанционных датчиков и сенсоров физической активности, диагностических приборов широкого пользования и др.) [22], отслеживающих состояния здоровья пациентов. Устанавливаемые электронные датчики производят миллионы транзакций в секунду, и необходимо надежное решение, способное трансформировать, сохранить и работать с ними в режиме реального времени. Безусловно, это под силу технологиям больших данных.

Стремительное развитие технологий мобильной беспроводной связи и программных приложений стимулировало разработку и распространение способов их инновационного применения для решения первоочередных задач медицины и здравоохранения. Это привело к формированию новой области электронной медицины, известной как «мобильное здравоохранение» (*Mobile Health, mHealth*) [23]. По данным

Международного союза электросвязи (МСЭ), на сегодняшний день число абонентов беспроводной связи в мире приближается к 5 миллиардам человек, и более 85% мирового населения находится в зоне покрытия сигнала коммерческих сетей беспроводной связи [24]. При этом большая часть данных, произведенных в период с 2012 по 2020 год, будет сгенерирована не людьми, а различного рода устройствами в ходе их взаимодействия друг с другом и сетями данных (сенсорами, смартфонами, устройствами радиочастотной идентификации (RFID), спутниковыми системами навигации и т.п.) [25].

Определение и сущность понятия «большие данные» медицине

Прежде чем перейти к исследованию понятия «Big Data» (Большие Данные) в контексте электронной медицины, рассмотрим, что же подразумевается под этим термином?

Анализ литературы показывает, что на сегодня термин «Big Data» не имеет строгого, общепринятого определения. Согласно [26] Big Data обозначает множество настолько объемных и сложных наборов данных, что их обработка традиционными инструментальными средствами за приемлемое для практики время довольно затруднительна. Согласно отчету McKinsey Institute [4], термин «Big Data» подразумевает набор данных, размер которых превосходит возможности типичных баз данных по сбору, хранению, управлению и анализу. В настоящее время большими считаются данные объемом от терабайта (1 терабайт = 1024 гигабайта) и более. Big Data обычно измеряются в терминах «терабайты, петабайты, эксабайты и т.д.» (1 эксабайт = 1 миллиард гигабайт). В отчете, представленном в августе 2012 года в Конгресс США, предлагается следующее определение: «Big Data» – это разнородные данные больших объемов с высокой скоростью изменения, для обработки которых необходимо применение инновационных методов и технологий обработки, хранения, передачи, управления и анализа информации» [10]. Еще одно определение этого термина приведено в [27]: «Big Data – это совокупность инструментов и методов обработки структурированных и неструктурированных данных огромных объемов из различных источников, подверженных постоянным обновлениям, в целях повышения качества принятия управленческих решений, создания новых продуктов и повышения конкурентоспособности».

Как следует из приведенных определений, сущность Big Data заключается в предоставлении возможности высокопроизводительного анализа быстро обновляющихся разнородных данных огромного объема посредством новых методов и технологий обработки информации в режиме реального времени. Целью обработки Big Data являются выявление взаимосвязи между всеми разнородными данными и выделение скрытых новых знаний, необходимых для поддержки принятия решений, создания новых видов продукции, повышения конкурентоспособности предприятий и т.п.

Рассмотрим определение, специфику и сущность понятия «Big Data» в контексте электронной медицины. Для этого, прежде всего, внесем ясность в смысл терминов «электронное здравоохранение» и «электронная медицина», к интерпретации которых на сегодня также нет однозначного подхода [28, 29], и уточним авторское понимание этих понятий.

Так, по определению, здравоохранение – это отрасль деятельности государства, целью которой являются организация и обеспечение доступного медицинского обслуживания населения, сохранение и повышение его уровня здоровья [30]. Иначе говоря, основными функциями отрасли являются организационно (административно)-управленческие, направленные на анализ данных, прогнозирование и эффективное планирование будущего развития на основе представительных статистических выборок, временных рядов и пр.

Медицина – система научных знаний и практических мер, объединяемых целью распознавания, лечения и предупреждения болезней, сохранения и укрепления здоровья и

трудоспособности людей, продления их жизни [30]. Медицина как наука об исцелении охватывает гигантскую систему постоянно растущих знаний и данных, объем которых затрудняет работу с ними. Поэтому применение современных технологий обработки медицинских знаний и данных, предоставляющих возможность своевременной поддержки принятия лечебно-диагностических решений, жизненно необходимо медицинским учреждениям. Основным приоритетом цифровизации должны быть задачи поддержки врачебной деятельности, непосредственно определяющие качество медицинских услуг, а не задачи учета статистических показателей [31]. Исходя из цели медицины и соглашаясь с той частью авторов, которые считают, что информатизация должна быть направлена на поддержку лечебно-диагностических решений, а информация нужна врачу, в первую очередь, для лечения пациента, а не занесения данных в учетную систему, мы в настоящей работе отдаем предпочтение термину «электронная медицина». Итак, мы будем обсуждать феномен Big Data в электронной (цифровой) медицине, рассматривая этот термин в более широком смысле, включающем также организационно-управленческие процессы охраны здоровья населения.

По определению, Big Data в медицине – это различные быстро пополняемые электронные медицинские данные колоссального объема, которыми невозможно управлять посредством традиционных инструментов и методов, программного и/или аппаратного обеспечения [32].

Основные показатели больших данных в медицине

Наиболее распространенным является раскрытие феномена Big Data через указание проблем, с которыми приходится сталкиваться на современном этапе технологического развития при обработке информации [33, 34]. К основным проблемам, характерным для Big Data, относятся большой объем (*Volume*), разнообразие (*Variety*) и высокая скорость изменения (*Velocity*) данных. Для некоторых сфер деятельности, в том числе и для медицины, исследователями и практиками введены еще две характеристики: достоверность (*Veracity*) и ценность (*Value*) больших данных [35]. Отметим тесную взаимосвязь этих проблем и приведем их интерпретацию в медицине.

Объем (*Volume*) медицинских данных. Непрерывное генерирование и накопление медицинской информации в ближайшие годы приведут к наличию невероятного объема данных [10]. В настоящее время медицинские данные включают в себя клинические данные, полученные врачами, личные медицинские записи пациентов (EHR), радиологические изображения (рентгеновские и маммологические снимки), МРТ (магнитно-резонансная томография), КТ (компьютерная томография) и т.п., данные лабораторий и аптек, страховые претензии и пр. Параллельно генерируется также огромное количество медицинской информации, не относящейся конкретно к пациенту. Это многочисленные медицинские публикации, научно-исследовательские отчеты, результаты исследований и разработок (R&D), социологические опросы и др. Новые типы медицинских данных, такие как многомасштабная 3D/4D визуализация, геномика, биометрические показания датчиков и т.п., также способствуют экспоненциальному росту этой информации. Объемы этих данных очень велики, результаты одного исследования могут занимать сотни мегабайт. Даже на уровне одной клиники ежегодный объем пополняемых данных, связанных с современными методами диагностики, может составлять терабайты или даже десятки терабайт. Например, только в США объем медицинской информации в 2011 году составил 150 экзабайт [36]. Сложность, а зачастую и невозможность как обработки и хранения Big Data, так и установления содержательных отношений и закономерностей между ними для извлечения полезной информации требуют применения новых технологических и алгоритмических решений.

Многообразие/разноформатность (Variety). Возможность одновременной обработки медицинских данных различного формата (структурированных, полуструктурированных и неструктурированных) является одной из специфических проблем медицины [36]. Более 70–80% медицинских данных являются неструктурированными, причем наблюдается устойчивая тенденция к значительному росту скорости этой информации относительно таковой для структурированных данных. Для принятия лечебно-диагностических решений появляется необходимость в интегрировании клинической информации и биологических данных, которые имеют различные форматы и генерируются разными гетерогенными источниками. Так, медицинские данные могут быть представлены в виде числовых значений в произвольных единицах измерения, изображений, текстов, рукописных заметок и рецептов врача, геномных и белковых последовательностей и пр. Совместное хранение, сравнение и конвертация разнородной информации требуют решения очень сложных задач, таких, как распознавание образов и изображений, сжатие информации и т.п. И здесь медицина далеко не всегда может воспользоваться готовыми решениями из других областей [37]. Необходимы технологии комплексного анализа разноформатных медицинских данных, позволяющие сделать их полезными для дальнейшей обработки.

Скорость (Velocity) производимых медицинских данных. Постоянный поток новых данных накапливается беспрецедентными темпами, т.е. наблюдаемый рост объема и разнообразия данных непосредственно связан со скоростью, с которой они генерируются. Специфика проблемы в медицине заключается в том, что скорость пополнения данных также накладывает ограничения на обработку Big Data в медицине. Так, информация от приборов, следящих за пациентами отделения интенсивной терапии, поступает непрерывно, т.е. в режиме реального времени, и требуются ее немедленная обработка и анализ для своевременной выработки предварительного диагноза [10, 36].

Достоверность (Veracity) данных. Этот признак отражает семантическую и синтаксическую определенность, качество, актуальность и надежность данных. Ряд проблем обеспечения достоверности данных специфичны именно для медицины, поскольку имеют дело с диагнозами, методами лечения, рецептами, процедурами и т.п. Так, результаты анализа больших данных должны быть безошибочными и надежными, поскольку от достоверности сделанного на основе анализа Big Data вывода может зависеть жизнь человека. С другой стороны, низкое качество медицинских данных, особенно неструктурированных, является одной из серьезных проблем: заполняемые часто с ошибками медицинские карточки, неверное истолкование и неточный цифровой ввод назначений врача в рецептах из-за плохого почерка относятся к разряду наиболее распространенных примеров. Достоверность данных может зависеть и от качества медицинских приборов и их сенсоров (например, провода могут оборваться и сигнал может оказаться слабым из-за того, что сенсор размещен не в том месте; на показатели могут повлиять и электрические помехи) [36, 38].

Слишком высокая цена ошибки является причиной недоверия медицинской общественности к результатам анализа больших данных.

Ценность (Value) накопленных данных. Этот показатель, представляющий интерес для различных заинтересованных сторон и лиц, принимающих решение, характеризует Big Data с позиций их полезности и привнесения определенной ценности для медицинского учреждения (МУ) и системы здравоохранения в целом (например, усовершенствования бизнес-процессов, определения бизнес-стратегий, оптимизации расходов и т.п.) [35].

В условиях стремительного роста объема данных будут меняться также модели расходов на здравоохранение и стоимость лечения. Хотя прибыль не является и не должна быть основным мотиватором, однако она жизненно важна для организаций здравоохранения для того, чтобы приобрести доступные инструменты, инфраструктуру и методы эффективного использования больших данных. В противном случае МУ (клиники,

больницы, службы медицинского ухода на дому, центры реабилитации и др.) рискуют потерять потенциально миллионы долларов в виде доходов и прибыли [37, 39].

Анализ характерных особенностей медицинской информации, с одной стороны, и специфических возможностей Big Data – с другой, позволяет сделать вывод о целесообразности применения последних для обработки огромного количества медицинских данных. Так, потенциал и суть Big Data в медицине, с нашей точки зрения, можно определить как: 1) предоставление последними новых технологий и инструментария для хранения, передачи, быстрой обработки и анализа непрерывно генерируемых разнородных и разноформатных данных, поступающих из разных источников; 2) выявление корреляций между различными, казалось бы, не связанными медицинскими показателями и факторами, воздействующими на них; 3) получение результатов обработки и анализа в виде информации (знаний), понятной для врача и пригодной для принятия обоснованных решений.

Направления приложения и потенциал Big Data в медицине

В течение последних десятилетий в мире наряду с тенденциями к повышению эффективности медицинской помощи наблюдается стремительный рост расходов на здравоохранение. Проблема актуализируется в связи с продолжающимся старением населения во многих развивающихся и развитых странах и быстрым ростом возрастной когорты граждан в возрасте старше 65 лет, широким распространением таких «болезней образа жизни», как ожирение и диабет, а для когорты пожилых – нарушения памяти, артрит и т.п. Однако практически исчерпанные возможности экстенсивного развития системы медицинской помощи путем наращивания ее объемов, соответственно и затрат, не имеют перспектив. В этой связи уже сегодня разрабатываются новые концептуальные подходы к решению этой проблемы, рассматривающие Большие Данные как дополнительный источник компенсации расходов, с одной стороны, и повышения качества медицинских услуг – с другой.

Роль Big Data-аналитики в медицине

Аналитика является инструментом или набором методик, преобразующих исходные данные в полезную информацию. Big Data в качестве аналитической основы новых ИТ-приложений в медицинской индустрии могут использоваться для поддержки принятия как лечебно-диагностических, так и организационно-управленческих решений. Так, например, эта информация может использоваться медицинскими организациями для поддержки принятия обоснованных стратегических и оперативных решений, оптимизации расходов, перспективного прогнозирования и создания дополнительной стоимости. С развитием технологий и инструментов аналитики Big Data медицинские организации могут легко управлять колоссальным количеством поступающих из различных источников разнородных и разноформатных цифровых данных. Здесь аналитика играет очень важную роль в отделении полезной информации от бесполезной [40]. Пациент-ориентированный подход и Big Data в качестве аналитической основы новых ИТ-решений в медицине имеют потенциал для преобразования почти каждого направления этой отрасли. Анализ литературы позволил выделить ряд направлений приложения Big Data в медицине, перспективных с позиций потенциала последних.

Big Data в поддержке принятия решений по менеджменту качества медицинского обслуживания

Растущая доступность медицинских данных позволяет с помощью технологий Big Data предложить более глубокое понимание взаимосвязей, лежащих в основе огромного количества информации из различных множеств данных, и преобразовать последние в новые знания. В результате анализа Big Data становится возможным выявить такие неожиданные взаимосвязи или закономерности, которые человек не в состоянии

обнаружить. Так, например, технологии Big Data позволяют одновременно обрабатывать базы историй болезней пациентов, геномных данных и отчетов медицинских исследований и в итоге выдавать полезную информацию для принятия наилучшего решения относительно лечения конкретного пациента [38].

В сущности термины «анализ Big Data» и «принятие решений» относятся к реализации процесса тщательного анализа огромного объема больших данных с целью выявления определенных тенденций и выработки обоснованных решений [41].

Big Data в поддержке принятия врачебных решений. По оценкам McKinsey, внедрение технологий анализа Big Data в медицину может сэкономить для здравоохранения США около 300 миллиардов долларов в год. Экономия произойдет за счет более своевременной и точной постановки диагноза, подбора эффективного лечения, сокращения расходов на исследования. Ожидается, что эффективность лечения будет повышена благодаря обработке всей доступной информации. В практику врачей войдет широкое использование систем поддержки принятия решений и экспертных систем нового поколения, позволяющих предоставить врачам невиданный ранее доступ к опыту коллег посредством анализа EHR пациентов как в различных географических срезах, так и в масштабе страны. Это, в свою очередь, позволит минимизировать субъективный человеческий фактор при принятии врачебных решений о стратегии лечения пациента.

Для подтверждения состоятельности этого прогноза достаточно отметить, что экспертная система (ЭС) Watson фирмы IBM в 2013 году на общих основаниях сдала экзамены и получила диплом врача, приобретя юридическое право лечить людей. В области онкологии ЭС уже демонстрирует блестящие результаты. Она способна изучить историю болезни пациента, записи и комментарии врачей, просмотреть последние исследования по данной теме и предлагать диагнозы на основе всех изученных данных. При этом IBM Watson детально анализирует данные, сопоставляет различные факторы и проводит аналогии. В качестве исходных данных в память суперкомпьютера загружено более 600 тысяч медицинских заключений и диагнозов, 2 миллиона страниц текстов, взятых из 42 медицинских журналов и результатов клинических испытаний в области онкологии. Watson может «проанализировать» 1,5 миллиона записей из историй болезни различных пациентов и, основываясь на данных из историй успешной борьбы с подобными заболеваниями (наилучшей практики), выявить наиболее подходящие методы лечения в каждом конкретном случае [42].

В настоящее время к одной из тенденций, приобретающих значимость, относится также поддержка клинических решений для управления здоровьем отдельного пациента с учетом совокупности данных населения в целом. И здесь на помощь может прийти уникальное преимущество технологий анализа Big Data, выраженное в их способности анализировать данные всего населения с учетом различных демографических показателей, а также сегментировать население по отдельным когортам. Так, сегодня становится реальностью с помощью технологий Big Data сравнить в рамках одной системы данные конкретных пациентов за прошлые годы с результатами текущих лабораторных исследований и одновременно связать эту информацию с демографическими данными, ретроспективной информацией о реакции на препараты и взаимодействиях лекарств пациентов любого возраста, пола и этнической принадлежности. Эти возможности Больших Данных позволяют строить как географические и социальные модели здоровья населения, так и предиктивные модели развития эпидемических вспышек, оказывая этим поддержку решениям по предупреждению последних [43].

Потенциал Big Data в стандартизации врачебных решений. Каждый пациент в своей практике не раз сталкивался с ситуацией, когда два разных врача совершенно по-разному интерпретировали результаты его анализов или одно и то же радиологическое изображение (медицинские данные) и ставили разные диагнозы. Анализ больших данных, т.е. накопленного статистического материала по определенным заболеваниям, и его

сравнение как с текущими данными конкретного пациента (результатами лабораторных анализов, рентгеновскими снимками и т.п.), так и с его EHR может использоваться для поддержки принятия врачебных решений. Минимизация субъективного человеческого фактора в процессе сопоставления разнородных данных пациента, его симптомов и огромного объема статистики отдельных болезней предоставляет врачу возможность выбора наилучшего варианта решения, что соответствует наиболее вероятному диагнозу, т.е. способствует стандартизации медицинских решений. В клинической практике нередко возникает необходимость в стандартизации медицинских решений. Так, например, при разборе врачебных ошибок, судебных разбирательствах, медицинской экспертизе, неправильной диагностике и лечении, выявлении эффективности назначенных медикаментов в каждом конкретном случае (заболевании). В этих ситуациях Big Data могут оказать поддержку в принятии обоснованных решений (выявлении истины) [44].

Codeйствие Big Data развитию и созданию персонализированной медицины. Исследования показывают, что в ближайшем будущем качество врачебных решений уже не может и не будет опираться на опыт и интуицию: конкурентные преимущества будут достигаться с учетом прогнозирования последствий принимаемых решений, актуализирующей необходимость создания принципиально новой модели организации медицинской помощи, так называемой персонализированной медицины (personalized medicine), базирующейся на концепции 4П. Концепция 4П-медицины интегрирует в себе понятия *персонализации* (индивидуального подхода к каждому пациенту с учетом его генетических особенностей), *предиктивности* (прогнозирования предрасположенности к развитию заболевания), *превентивности* (предотвращения или снижения риска развития болезни), *партисипативности* (мотивированного участия пациента в профилактике возможных заболеваний и их лечении) [45]. Сущность персонализированной медицины заключается в индивидуализации лекарственной терапии в соответствии с персональными данными и генотипом конкретного пациента. Как новая парадигма в здравоохранении, персонализированная медицина предполагает раннее (доклиническое) выявление заболеваний на этапе прогнозирования предрасположенности и последующих превентивных мероприятий [45, 46].

Технологии Big Data могут сыграть существенную роль в реализации персонализированной медицины, ориентированной на выбор правильного лечения. Так, системы поддержки принятия решений, базирующиеся на Big Data, могут на основе обработки и анализа гигантских объемов генетической информации пациентов прогнозировать их реакцию на те или иные лекарственные препараты, назначать абсолютно уникальные лекарственные средства в индивидуальных дозах [46].

Процесс персонализации медицины, т.е. перехода от врач-ориентированного подхода к оказанию медицинских услуг к пациент-ориентированному на основе индивидуализации медицинской помощи, уже происходит [47]. Так, более 20% новых лекарств, продуктов питания и медикаментов, одобренных в 2014 году в США Коалицией персонализированной медицины (основанной в 2004 году), было продукцией персонализированной медицины [48]. В некоторых американских клиниках база данных генотипов интегрирована с системой лекарственных назначений EHR и экспертной системой для персонификации лекарственной терапии. Такие системы четвертого поколения, по классификации Gartner «коллеги врача» с продвинутыми аналитическими возможностями, уже начали распространяться в мире [49].

Потенциал Big Data в поддержке удаленного мониторинга здоровья пациентов. Технологии обработки Больших Данных могут способствовать развитию персональной и превентивной медицины, основанной на удаленном мониторинге пациентов. Это, в свою очередь, будет содействовать повышению качества жизни пациентов путем доступа последних к медицинским услугам в любое время и в любом месте, своевременного

обнаружения и предотвращения неожиданных осложнений в состоянии пациента, поддержки (автоматического напоминания) проведения пациентом самостоятельных лечебно-профилактических процедур.

Интегрирование EHR пациентов (медицинские тесты, лабораторные анализы, предписанные лекарства, информация о побочных эффектах, противопоказаниях, неструктурированные тексты и т.п.) с подключенными к Интернет носимыми smart-устройствами, контролирующими жизненно важные показатели здоровья (циклы сна, частоту сердечных сокращений, давление и т.д.), позволяет врачам наблюдать за больным в режиме реального времени. Консолидация непрерывно генерируемых в процессе отслеживания состояния пациента медицинской информации с данными EHR и одновременный анализ огромного объема разнородных данных в масштабе реального времени предоставляют профессионалу комплексную связную картину общего состояния здоровья пациента. Это, в свою очередь, будет способствовать поддержке принятия обоснованных решений по дистанционному диагностированию больного в режиме реального времени. В то же время дистанционный мониторинг здоровья пациентов позволит сократить затраты за счет отпадания необходимости в лабораторных исследованиях, госпитализации, своевременной коррекции тактики лечения.

Будущие приложения больших данных в режиме реального времени, такие как раннее обнаружение инфекции и применение превентивных мер, могут уменьшить заболеваемость и смертность пациентов и даже предотвратить вспышки эпидемий [10].

Согласно прогнозам [36, 50], возможность выработки в режиме реального времени аналитики меняющихся данных большого объема по всем направлениям медицины может произвести революцию в этой сфере.

Возможности Big Data в содействии развитию доказательной медицины. При принятии решения о выборе тактики лечения врачи традиционно используют свои суждения, но в последние несколько лет был сделан шаг к доказательной медицине. Доказательная медицина (*Evidence-based medicine*) – это концептуальный подход к медицинской практике, предполагающий принятие решений о тактике лечения больного на основе достоверных результатов, полученных в ходе множества клинических исследований [51]. Иначе говоря, доказательная медицина подразумевает принятие наилучшего решения о тактике лечения на основе систематического обзора данных клинических испытаний.

Большие данные имеют огромный потенциал для создания доказательной базы по поддержке принятия врачебных решений. Так, доказательная медицина базируется на выводах, полученных в ходе рандомизированных контролируемых испытаний нового метода лечения, выполненных на ограниченном количестве пациентов. Однако в реальности могут иметь место достаточно редкие нюансы, отрицательно воздействующие на конечные результаты, которые невозможно выявить в процессе исследований на малых выборках. Объединение множества отдельных наборов данных в алгоритмах Big Data может предоставить самые надежные доказательства в выборе тактики лечения [52, 53].

Потенциал Big Data в содействии доступу к новым знаниям. Оцифровка медицинской литературы значительно расширяет также возможности доступа врачей к результатам новых исследований и технологий лечения, внедренных в клиническую практику. Однако стремительный рост количества медицинских открытий и скорость их появления в различных источниках по всему миру значительно превосходят физические возможности клиницистов по ознакомлению со всеми достижениями даже в разрезе отдельных заболеваний. Так, например, в мире ежегодно проводится около 170 тысяч клинических исследований препаратов для борьбы с онкологическими заболеваниями, однако выход к информации об их результатах достаточно ограничен [40, 54]. Большие данные могут облегчить выход к последним достижениям в мировой клинической практике и тем самым расширить возможности врача для быстрого получения новых знаний.

Автоматизация обследования и обработки рынка медицинских знаний позволит довести информацию о медицинских инновациях, новых лекарственных препаратах, созданных ведущими учеными в любой точке мира, до каждого заинтересованного лица [55]. В практике врача Big Data могут позволить ему пополнить или подтвердить свои знания, изучив в реальном времени данные, полученные от множества специалистов, занимающихся лечением пациентов со схожими заболеваниями. Улучшение поиска знаний путем доступа к высокопроизводительным и очень точным базам данных медицинских карт пациентов позволит предложить профилактическое лечение, выявить успешные шаблоны лечения и снизить число медицинских ошибок [56].

Возможности Big Data для поддержки пациентов в принятии ими решений по управлению собственным здоровьем. Как было отмечено выше, цифровизация медицинской отрасли способствовала трансформации многих ее аспектов. Наиболее заметные изменения произошли во взаимоотношениях пациентов и поставщиков медицинских услуг: врачей, медицинских центров, лабораторий и т.п. Появление многочисленных и разнофункциональных носимых устройств и сервисов оказало серьезное воздействие на представления пациентов о принципах оказания медицинских услуг. Согласно результатам исследования консалтинговой компании *PricewaterhouseCoopers* [57], приоритеты пациентов несколько поменялись, и теперь первоочередной необходимостью для них является легкий доступ к медицинским услугам и результатам лечения, с одной стороны, и потребность стать активным участником процесса лечения – с другой. Концепция «ответственного» пациента, стремящегося как можно больше знать о своем здоровье и имеющего возможность самостоятельно принимать решения посредством доступа к медицинским знаниям из альтернативных источников, становится все более популярной в мире. Это, в свою очередь, стимулирует спрос на медицинский контент и многочисленные сервисы, ориентированные на пациента. Усилению тренда способствует также политика многих государств, направленная на повышение здоровья нации и пропаганду правильного (здорового) образа жизни [58].

Согласно прогнозам, будущая медицина предполагает личное участие или вовлеченность пациента, которая станет возможной в результате появления новых сервисов и инструментов контроля своего состояния. В сложившейся ситуации целесообразно предоставление пациентам (физическим лицам) возможности квалифицированного отслеживания и анализа состояния здоровья, доступа к информации о методах лечения касательно их конкретных симптомов, позволяющих им принять обоснованные решения по более эффективному управлению собственным здоровьем.

Big Data могут помочь в реформировании взаимоотношений пациентов и поставщиков медицинских услуг в направлении повышения вовлеченности пациентов в процесс лечения. В соответствии с действующей на сегодня моделью медицинские записи пациентов (EHR) находятся в распоряжении медицинского учреждения (МУ). В дальнейшем предполагается передача персональных медицинских данных (ПМД) в распоряжение пациентов. Big Data позволят дополнить ПМД пациентов в EHR их личными данными из различных веб-сайтов и социальных сетей. В целом, дальнейшее развитие решений в области электронной медицины ориентировано на создание платформы для передачи и обмена информацией, взаимодействия и сотрудничества между МУ, врачами-специалистами, медицинскими работниками и пациентами [59].

Потенциал Big Data в формировании клиент-ориентированной медицины. В настоящее время поставщики медицинских услуг стремятся к внедрению эффективной стратегии совершенствования процессов удержания пациентов (клиентов) путем использования внутренних, внешних и аналитических источников для обеспечения единого представления о клиенте. Появление новых моделей взаимоотношений врачей и пациентов вследствие цифровизации медицины вынудило многие медицинские организации принять

более клиент-ориентированный подход. При этом большинство организаций имеют представление о клиенте, исходя только из внутренних источников данных. Это создает трудности при попытке формирования всеобъемлющего представления о клиенте и требует привлечения также внешних данных.

Big Data позволяют дополнить персональные медицинские данные пациентов в EHR информацией, полученной из внешних источников, предоставляя широкие возможности для выявления факторов риска или так называемых факторов образа жизни. Так, в современном цифровом обществе каждый человек генерирует огромное количество информации в Интернет посредством личных данных на различных веб-сайтах, профилей в социальных сетях, использования кредитных карт и т.п. Эта информация позволяет выявить факторы образа жизни человека (например, материальное состояние, уровень образования, привычки, интересы, возможные заболевания и др.) без необходимости интервьюирования последнего. Совместное использование «сетевой» информации и данных пациента из EHR предоставляет уникальную возможность интегрирования традиционных медицинских моделей с социальными детерминантами здоровья пациента и формирования клиенториентированной медицины [60].

Потенциал Big Data в управлении лечением пациентов с хроническими заболеваниями. За последние несколько десятилетий практически во всем мире наблюдается существенный рост пациентов с хроническими заболеваниями. Значительно увеличилось также показатели коморбидности (comorbidity), т.е. одновременного наличия у пациента двух или более хронических заболеваний [61]. Это требует наблюдения таких пациентов не только у своего основного врача, но и у других специалистов. Однако разрозненность как медицинских учреждений, оказывающих разные услуги одним и тем же пациентам, так и методов лечения снижает способность системы здравоохранения предоставлять качественные медицинские услуги, приводит к ненужному дублированию анализов и лечения. В этой ситуации наибольшему риску подвергаются пациенты, нуждающиеся в комплексном медицинском обслуживании.

Например, пациент с застойной сердечной недостаточностью, диабетом и хроническим заболеванием легких нуждается в комплексном лечении, которое обеспечивало бы координацию и сбалансированность различных видов терапии. Однако на практике такая координация осуществляется редко. Вместо этого пациенты зачастую получают противоречащие друг другу рекомендации от различных врачей.

На сегодня пациенты, страдающие хроническими заболеваниями, являются крупнейшими потребителями ресурсов здравоохранения. Например, на их долю в здравоохранении США приходится 75% расходов этой сферы [62]. Big Data могут сыграть решающую роль в снижении расходов на лечение пациентов с хроническими заболеваниями. Так, с помощью методов интеллектуального анализа данных можно найти максимально похожие когорты пациентов, воспринимаемых в качестве кандидатов для профилактического вмешательства в группу высокого риска (Highriskgroup, HRG) [63]. Это может способствовать снижению расходов здравоохранения за счет: 1) улучшения координации ухода за пациентом; 2) обнаружения дублирования данных и уточнения лечебной терапии; 3) снижения повторных госпитализаций посредством выявления факторов образа жизни (наследственных, социально-экономических, профессиональных, экологических, поведенческих и др.), увеличивающих риск отклонений в состоянии здоровья; 4) формирования клиенториентированной медицины.

Возможности Big Data в поддержке принятия организационно-управленческих решений

Анализ существующей врачебной практики с использованием технологий Больших Данных может способствовать уменьшению затрат на работу больниц, устранить множество

злоупотреблений и лишних расходов в государственных системах здравоохранения. Так, МУ, используя огромные объемы разнообразных данных о пациентах, могут снизить затраты за счет оптимизации работы клиник: прогнозирования ожидаемого потока пациентов, сокращения очередей, оптимального ресурсного обеспечения, оценки производительности медицинского персонала и сбалансированности его нагрузки, повышения удовлетворенности пациентов качеством медицинских услуг и т.п. [64].

С помощью технологий Big Data можно совершенствовать и регулировать систему ценообразования и оплаты. Так, например, в США в соответствии с новыми федеральными законами в МУ, охваченных программой Medicare, производительность медицинского персонала непосредственно связана с платежами и штрафами за оказанные медицинские услуги, измеряемые такими показателями, как расходы пациентов, степень их удовлетворенности, повторная госпитализация и смертность. Ожидается, что к 2018 году примерно 90% медицинских услуг в США будет оплачиваться по программе Medicare [65], т.е. в зависимости от результатов лечения. Перейти к реализации системы оплаты и ценообразования, основанной на совместной оценке производительности медперсонала, расходов на медикаменты и затраченный труд, реальные финансовые возможности населения и потребность в услугах, возможно только на базе систем, работающих с большими данными.

МУ стремятся улучшить качество лечения пациентов при одновременном снижении затрат. Для повышения производительности МУ стремятся использовать и анализировать большие объемы разнородных внутренних и внешних данных о пациенте: клинические данные, информацию из EHR, данные мониторинга с медицинских приборов, непрерывно поступающие данные следящих систем (устройств образа жизни) и т.п. Это позволяет им повысить качество и разнообразие медицинских услуг, оказываемых гетерогенным группам пациентов. Растущая вовлеченность пациентов в заботу о собственном здоровье и спрос на дистанционные медицинские услуги позволяют МУ договориться о лучших условиях оплаты со стороны плательщиков (пациентов, страховых компаний и регулятора) [66]. Технологии Big Data могут сыграть также важную роль в извлечении и предоставлении ценной информации из колоссального количества данных поставщикам медицинских услуг и ЛПР для разработки стратегий, планов и принятия важных управленческих решений.

Преимущества Big Data и соответствующей аналитики в поддержке принятия медицинских решений

Обобщенно преимущества Big Data в медицине можно выразить в их содействии повышению качества медицинских услуг и предоставлении дополнительных источников сокращения затрат посредством:

- возможности анализа совокупности медицинских данных и поддержки принятия лечебно-диагностических решений как в разрезе всего населения и его географических срезов, так и в разрезе различных когорт и отдельных пациентов;
- развития доказательной медицины путем поддержки принятия более обоснованных стратегий лечения пациентов;
- развития персонализированной медицины, базирующейся на индивидуальном подходе к профилактике, диагностике и лечению генетически обусловленных заболеваний;
- автоматической консолидации данных пациента из сложных гетерогенных источников и их совместного использования в режиме реального времени;
- идентификации максимально схожих когорт пациентов, нуждающихся в комплексном медицинском обслуживании, и поддержке принятия скоординированных решений по их лечению;

- принятия организационно-управленческих решений, направленных на сокращение расходов, модернизацию системы оплаты и ценообразования.

Вызовы в использовании Big Data-аналитики в медицине

Наряду с потенциалом Big Data и растущими тенденциями практического использования последних существуют определенные вызовы на пути их широкого внедрения в медицину. Так, получению эффективной Big Data-аналитики препятствуют:

- непрерывный рост объема медицинской информации, лишь малая часть которой используется во всем мире;

- преобладание в системе EHR неструктурированных, неполных, неточных и случайных медицинских данных, влияющее на достоверность информации (неточный цифровой ввод назначений врача в рецептах из-за плохого почерка, низкое качество информации, полученной из удаленных источников, отсутствие информации в одном или нескольких полях EHR, увеличение мусора и пр.) [36, 67, 68];

- сложность точной идентификации личности пациента при сопоставлении данных, полученных из разнообразных источников (EHR, неклинических данных из социальных медиа, веб-сайтов и др.) для формирования целостной картины о состоянии здоровья пациента [38, 69, 70];

- растущая потребность в обработке непрерывно генерируемых данных пациента и выработке аналитики в масштабе реального времени, требующая выполнения сложных процессов интеграции, которые сегодня недоступны для многих клиник как в силу высокой стоимости технологий Больших Данных, так и отсутствия соответствующих специалистов в ИТ-отделах МУ. Последние должны также гарантировать безопасность, достоверность и высокую доступность данных из различных источников [71];

- потенциальные угрозы конфиденциальности и информационной безопасности персональных медицинских данных, связанные с опасностью нецелевого использования последних. Еще одна серьезная угроза безопасности персональных медицинских данных появилась после широкой доступности генетической информации, однозначно указывающей на конкретного пациента. Поскольку эту информацию практически невозможно анонимизировать, а деидентифицированные геномные данные легко восстановить, то в этом случае вопросы конфиденциальности должны быть решены на законодательном уровне [37, 72, 73];

- недостаточный объем накопленных информационных ресурсов вследствие отсутствия EHR и EMR, хранилищ различных медицинских данных во многих странах;

- нехватка специалистов с глубокими аналитическими знаниями, работающими с Big Data (Data Scientists), которая оценивается к 2018 году в 140–190 тысяч человек [4, 74];

- высокая стоимость и сложность адаптации инструментов, методов и алгоритмов анализа Big Data ограничивают круг медицинских организаций, имеющих возможность внедрить данные технологии. Дополнительным барьером на пути внедрения Big Data является необходимость в постоянном финансировании поддержки базы данных в актуальном состоянии. Например, компания IBM вложила в разработку Watson примерно 3 миллиарда долларов при том, что суперкомпьютер прошел «обучение» лишь по нескольким заболеваниям [75].

Технологии и методы обработки и анализа Big Data

В соответствии с [76, 77] под понятием «технологии Big Data» подразумеваются подходы, комплексы инструментальных средств (платформ), методы и техники (способы и процедуры) обработки огромных объемов разноформатных и многообразных данных для получения воспринимаемых человеком результатов.

Методы и техники анализа Big Data. Для объединения, хранения, манипулирования, анализа и визуализации Больших Данных разработан и адаптирован достаточно широкий спектр методов и техник. Последние заимствованы из многих областей, включая статистику, информатику, искусственный интеллект, прикладную математику, лингвистику и т.п. Не претендуя на полноту, перечислим некоторые из методов и техник анализа Big Data, выделенных в работах [4, 43, 78]: методы класса Data Mining, краудсорсинг, интеграция разнородных данных, визуализация, статистический анализ, прогнозная аналитика, предиктивное моделирование, машинное обучение, обработка естественного языка, искусственные нейронные сети, обучение ассоциативным правилам, классификация, кластерный анализ, регрессионный анализ; искусственные нейронные сети, сетевой анализ, оптимизация, генетические алгоритмы; распознавание образов и т.д.

Инструментальные средства обработки и анализа Big Data. Основными требованиями при разработке программного обеспечения систем для работы с Большими Данными являются включение в их состав средств параллельной обработки и распределенного хранения данных, переход от моделей, используемых в традиционных реляционных СУБД, к новым моделям, поддерживающим постоянное изменение структуры данных, горизонтальное масштабирование посредством кластеров из недорогих серверов умеренной производительности (commodity computers) и т.п. Так, например, как развитие зрелых реляционных баз данных SQL появился новый класс систем NoSQL (Not Only SQL), таких как HBase, Cassandra, MongoDB, Neo4j, Riak и др. [79].

Наиболее распространенными системами обработки Big Data являются платформа Hadoop и парадигма распределенных вычислений MapReduce. Hadoop обладает потенциалом для обработки очень больших объемов разнородных входных данных задачи, главным образом путем распределения последних по многочисленным серверам (узлам), в каждом из которых производится обработка определенного набора данных, а затем результаты подзадач сводятся в конечный (итоговый) результат. Это позволяет использовать Hadoop для реализации поисковых и контекстных механизмов высоконагруженных сайтов – Yahoo, Facebook, Amazon и др. Открытие со стороны Yahoo кода Hadoop сообществу способствовало появлению в ИТ-индустрии целого направления по созданию продуктов на базе Hadoop, разработчиками которых являются как известные мировые компании, так и стартапы.

За последние годы разработано много программных продуктов, облегчающих работу с Big Data, в том числе SAP HANA, высокопроизводительная NewSQL платформа для хранения и обработки данных, приложения по предварительной обработке потоков информации (Marshalling), решения для визуализации и самостоятельного анализа (Advanced Analytics with Self-Service Delivery) и другие. Наиболее популярными языками программирования являются Python, Java и Scala.

Доступность технологий Big Data значительно возросла благодаря облачным вычислениям, обеспечивающим новый уровень скорости обработки данных [80]. На рынке представлено множество отдельных облачных сервисов по аналитике Big Data, а также облачных платформ для развертывания приложений. С их помощью аналитику на базе Big Data можно использовать «здесь и сейчас», просто подключившись к Интернету и выбрав подходящий сервис.

Следует отметить, что в процессе внедрения разработанных на сегодня программных продуктов в медицину пользователи сталкиваются с рядом существенных недостатков, связанных со спецификой отрасли, которые требуют своего решения. Далее, предлагаемые платформы/инструменты требуют наличия специальных навыков программирования, которыми типичный конечный пользователь в здравоохранении не может обладать. Поскольку технологии анализа Big Data начали использоваться в медицине недавно, то

предстоит еще решить вопросы управления данными, неприкосновенность частной жизни, безопасности, стандартизации и т.п.

Перспективы развития Big Data в медицине

Согласно ABI Research в настоящее время процессы анализа Big Data в медицине находятся в начальной стадии. Однако по прогнозным расчетам к 2019 году в результате стремительного роста объема непрерывно генерируемых данных рынок Big Data-анализа составит примерно \$52 миллиарда [71]. Аналитики TechNavio прогнозируют увеличение расходов на глобальном рынке Big Data в отрасли медицины в течение 2014-2019 гг. со среднегодовым приростом в 42% [81].

Ниже приведены наиболее перспективные направления развития и внедрения Big Data в медицину [82–86]:

- развитие мобильной медицины, расширяющей возможности удаленной диагностики и непрерывного мониторинга состояния здоровья отдельных категорий пациентов, актуализирующей спрос на потоковую обработку информации путем выработки аналитики на базе комплекса данных пациента в режиме реального времени;
- содействие технологий в поддержке принятия обоснованных лечебно-диагностических решений на основе сравнительного анализа данных пациента с базами историй болезней по практике лечения схожих заболеваний и выявления наиболее подходящего метода лечения в каждом конкретном случае;
 - расширение внедрения Big Data в реализации персонализированной медицины;
 - расширение фундаментальных исследований в сфере разработки оптимальных способов обработки, анализа и извлечения полезных знаний из больших объемов медицинских данных различной природы;
 - подготовка кадров для работы с большими медицинскими данными;
 - развитие и адаптация инструментов для работы с Big Data для медицинской отрасли: тренды на развитие мобильных приложений, облачной инфраструктуры, визуализации данных и т.п.
- Big Data-решения для обработки медицинских данных, сгенерированных «умными» вещами в рамках экспоненциально растущего направления – Интернета вещей.

Заключение

Проведенный анализ состояния Big Data в медицине показывает, что в настоящее время этот сегмент рынка находится в начальной стадии развития. Однако всё растущая цифровизация медицины, превращение Интернет в один из мощных источников медицинской информации и инфраструктуры для быстрой, надежной и безопасной передачи данных, развитие социальных сетей, мобильных и беспроводных технологий доступа к медицинской информации способствовали трансформации многих аспектов медицинской отрасли и обусловили появление новых трендов. Здесь можно отметить ряд таких новых тенденций, как: 1) развитие персонифицированной медицины, предполагающей эффективно организованный доступ к любой совокупности медицинских записей и первичных результатов исследований пациента; 2) изменение роли пациентов, повышение их вовлеченности в заботу о собственном здоровье и возможность его самостоятельного мониторинга; 3) изменение способов взаимодействия врачей и пациентов, доступа последних к медицинским услугам, результатам лечения и медицинской информации; 4) предоставление медицинских услуг в режиме реального времени; 5) развитие персонализированной медицины, базирующейся при назначении терапии на индивидуальную чувствительность пациента к лекарствам, и т.п. Все эти трансформации происходят на фоне одной общей цели, сводящейся к повышению

эффективности и гибкости системы здравоохранения посредством улучшения качества медицинских услуг при одновременном снижении затрат.

Big Data могут оказать существенное воздействие на здравоохранение. Колоссальный объем данных, генерируемых в настоящее время медицинскими учреждениями, в ближайшие годы будет расти с еще большей скоростью. Это неизбежно приведет к повышению спроса на анализ больших данных. Извлечение из этих данных полезной информации может принести большую пользу в разработке новых методов и технологий лечения, идентификации болезней и их профилактике, уменьшении количества медицинских ошибок, обеспечении общей безопасности здоровья населения страны, совершенствовании системы здравоохранения в направлении реализации доказательной и персонализированной медицины. Это, в конечном счете, приведет к улучшению качества лечения и уменьшению числа больных и, соответственно, расходов затрат медицинских учреждений.

Следует отметить, что хотя потенциально Big Data в медицине могут способствовать решению многих проблем, однако на сегодня очень мало реальных примеров успешного применения этой новой технологии в медицине. Аналитики и практики видят причину этого в недостаточной адаптации имеющихся на рынке платформ Big Data к специфическим особенностям медицинской отрасли, требованиям к конфиденциальности и безопасности медицинских данных, стандартизации врачебных решений и т.п. Внедрению Big Data в медицину мешают также такие факторы, как недоверие врачей и неуверенность ряда ИТ-компаний в эффективности данных технологий. Это означает, что пути к использованию Big Data в медицине проходят через решение указанных проблем и совершенствование этих технологий. Динамика развития технологий и сервисов Больших Данных и прогнозы аналитиков усиливают уверенность в том, что в ближайшем будущем будет наблюдаться значительный рост внедрения Big Data-решений в медицину. Более того, по мнению последних дальнейшая цифровизация неизбежно приведет медицину к революционным изменениям, последствия которой окажутся намного масштабнее.

Литература

1. <https://ru.wikipedia.org/wiki>
2. The Australian public service big data strategy. Canberra: Commonwealth of Australia 2013 www.aiaa.com.au/documents/policy-submissions/policies-and-submissions/2014/pdf
3. Рост объема информации – реалии цифровой вселенной. www.tssonline.ru/articles2/fix-corp/rost-obema-informatsii-realii-tsifrovoy-vselennoy
4. James Manyika, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, Angela Hung Bayers. Big data: The next frontier for innovation, competition, and productivity. Analyst report, McKinsey Global Institute, May 2011. www.mckinsey.com
5. Raghupathi W. Data Mining in Health Care /Healthcare Informatics: Improving Efficiency and Productivity. Edited by Kudyba S. Taylor & Francis; 2010:211-223.
6. Sun, J., Reddy, C. Big data analytics for healthcare. International Conference on Data Mining. Austin, TX., 2013. www.siam.org/meetings/sdm13/sun.pdf
7. Manchini M. Exploiting Big Data for improving healthcare servuces// Journal of e-Learning and Knowledge Society, 2014, v.10, n.2, pp.23–33.
8. Roski J, Bo-Linn GW, Andrews TA. Creating value in health care through big data: opportunities and policy implications. Health Affairs, 2014 Jul.33(7):1115–1122. www.ncbi.nlm.nih.gov/pubmed/25006136
9. Datta M. How Big Data Will Lower Costs and Advance Personalized Medicine. GEN Exclusives, 2013. www.genengnews.com/insight-and-intelligence/how-big-data-will-lower-costs-and-advance-personalized-medicine/77899962/

10. Transforming Health Care through Big Data. Strategies for leveraging big data in the health care industry. Institute for Health Technology Transformation, 2013. http://c4fd63cb482ce6861463-bc6183f1c18e748a49b87a25911a05/iHT2_BigData_2013.pdf
11. Iain Buchan, Chris Bishop. A Unified Modelling Approach to Data-Intensive Healthcare. The fourth paradigm. 2009, pp.91–97.
12. Powell J., Buchan I. Electronic health records should support clinical research // Journal of Medical Internet Research, 2005, vol.7, no.1, p.e4, doi: 10.2196/jmir.7.1.e4.
13. Зингерман Б. Электронная медицинская карта и принципы ее организации. www.osp.ru/medit/blogs/bz/bz_109.html
14. Kashmir H. How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did // Forbes, 2012. <http://ghr.nlm.nih.gov/glossary=personalizedmedicine>
15. Personalized medicine. <http://ghr.nlm.nih.gov/glossary=personalizedmedicine>
16. Jain K.K. Personalized medicine// Advances in Clinical Chemistry, 2008; №4, pp.548–558.
17. Gillam M., Feied C, Handler J., Moody E., Shneiderman B., Plaisant C., Smith M., Dickason J. The Healthcare Singularity and the Age of Semantic Medicine. The fourth paradigm. 2009, pp.57–64
18. Davenport T., Glaser J., Just-in-time delivery comes to knowledge management, Harvard Business Review, , 2002, vol.80, no.7, pp.107–111, 126doi: 10.1225/R0207H.
19. Manhattan research. Social Media&Health 2.0. <http://manhattanresearch.com/Research-Topics/Consumer/HealthSocial-Media>.
20. Эванс Дэйв. Интернет вещей: как изменится вся наша жизнь на очередном этапе развития Сети. www.cisco.com/web/RU/news/releases/txt/2011/062711d.html
21. Internet-of-Everything makes everything possible... <http://www.internet-of-everything.no/>,
22. Gawande A. Annals of Health Care. NewYorker, 2012. www.newyorker.com/reporting/2012/08/13/120813fa_fact_gawande
23. mHealth: New horizons for health through mobile technologies.WHO. Global Observatory for eHealth series, Vol. 3, 112p. http://www.who.int/goe/publications/goe_mhealth_web.pdf.
24. The world in 2010: ICT facts and figures. Geneva, International Telecommunications Union, 2010 ([http:// www.itu.int/ITU-D/ict/material/FactsFigures2010.pdf](http://www.itu.int/ITU-D/ict/material/FactsFigures2010.pdf))
25. White T. Hadoop: The Definitive Guide. OReilly Media, 3rd Edition, 2012. http://cdn.oreilly.com/oreilly/booksamplers/9781449311520_sampler.pdf
26. http://en.wikipedia.org/wiki/Big_data
27. Knowledge management in organizations // Proceedings of 9th International Conference, КМО 2014, Santiago, Chile, September 2-5, 2014. <http://www.springer.com/gp/book>
28. What Is The Difference Between Telemedicine, Telecare and Telehealth? <http://evisit.com/what-is-the-difference-between-telemedicine-telecare-and-telehealth/>
29. Telehealth, Telecare and Telemedicine...What's the Difference? www.globalmed.com/additional-resources/telehealth-telecare-and-telemedicine.php
30. <https://ru.wikipedia.org/wiki>
31. Лагутин Ю. Врачи заинтересованы в экспертных базах знаний. www.computerra.ru/cio/497
32. Hesla L. Particle physics tames big data// Symmetry magazine, 2012. www.symmetrymagazine.org/article/august-2012/particle-physics
33. Laney D. 3D Data Management: Controlling Data Volume, Velocity and Variety // Application Delivery Strategies. META Group. 2001. <http://blogs.gartner.com/douglaney/files/2012/01/ad949-3D-Data-Management-Data-Volume-Velocity-and-Variety.pdf>
34. Савельев А.И. Проблемы применения законодательства о персональных данных в эпоху «больших данных» (BIG DATA). www.hse.ru/pubs/share/direct/document/150345956.
35. Herman R, Williams P. Big Data in healthcare: what is it used for? <http://ro.ecu.edu.au/cgi/viewcontent.cgi?article=1021&context=aeis>

36. Feldman B., Martin E., Skotnes, T. Big data in healthcare: Hype and hope. 2012. www.west-info.eu/files/big-data-in-healthcare.pdf
37. Mattison J. How to Navigate Big Data in Healthcare. www.cio.com/article/2851986/healthcare/how-to-navigate-big-data-in-healthcare.html
38. Panahiazar M., Taslimitehrani V., Jadhav A., Pathak J. Empowering personalized medicine with Big Data and Semantic Web Technology // Proc IEEE Int Conf Big Data. 2014, pp.790–795. www.ncbi.nlm.nih.gov/pmc/articles/PMC4333680/.
39. LaValle S., Lesser E., Shockley R., Hopkins MS., Kruschwitz N.: Big data, analytics and the path from insights to value. MIT Sloan Management Review, 2011, vol.52, no.2. www.ibm.com/smarterplanet/global/files/in_idea_smarter_computing_to_big_data.pdf
40. Raghupathi W., Raghupathi V. Big Data Analytics in Healthcare: Promise and Potential// Health Inform. Science and Systems, 2014, 2:3. www.hissjournal.com/content/2/1/3
41. How Data Analytics can help in Decision Making in Healthcare. A Saviance Technologies Whitepaper. <http://sharinghealthcaresolutions.covidien.com/how-do-you-use-big-data-drive>
42. IBM's Watson Gets Its First Piece Of Business In Healthcare www.forbes.com/sites/bruceupbin/2013/02/08/ibms-watson-gets-its-first-piece-of-bus
43. Крылов В.В., Крылов С.В. Большие данные и их приложения в электроэнергетике. Изд. Нобель-пресс, 2014, 166 с.
44. Hoffman S. Medical big data and big data quality problems. <http://insurancejournal.org/wp-content/uploads/2015/03/Hoffman.pdf>
45. Hood L. Systems biology and p4 medicine: past, present, and future // Rambam Maimonides Med. J. 2013, vol. 4 (2). p. e0012. doi: 10.5041/RMMJ.10112
46. Sacchi L., Lanzola G., Viani N., Quaglini S., Personalization and Patient Involvement in Decision Support Systems: Current Trends // IMIA Yearbook of Medical Informatics, 2015, vol. 10 (1), pp.106–118
47. Horvitz E. From Data to Predictions and Decisions: Enabling Evidence-Based Healthcare. Computing Community Consortium. 2010. www.cra.org/ccc/initiatives
48. PMC Analysis. 2015. http://www.personalizedmedicinecoalition.org/News/Press_Releases/
49. Поколения МИС по Gartner, 2013. <https://ru-ru.facebook.com/OLTechnologies/posts/8162>
50. Basel Kayyali, David Knott, and Steve Van Kuiken. The Big-Data revolution in Healthcare: Accelerating value and innovation. Full report. 2013. www.mckinsey.com/insights/health_systems_and_services/big-data_in_health_care
51. Evidence Based Medicine: A new paradigm for the patient. Evidence Based Medicine Working Group, 1993 <http://jama.jamanetwork.com/article.aspx?articleid=404143/>.
52. Murdoch T.B.; Detsky A.S. The Inevitable Application of Big Data to Health Care// Journal of the American Medical Association. 2013; 309(13):1351–1352. <http://jama.jamanetwork.com/article.aspx?articleid=1674245>
53. Discovery with Data: Leveraging Statistics with Computer Science to Transform Science and Society. www.amstat.org/policy/pdfs/BigDataStatisticsJune2014.pdf
54. Alper B.S., Hand J.A., Elliott S.G., Kinkade S., Hauan M.J., Onion D.K., Sklar B. M. How much effort is needed to keep up with the literature relevant for primary care? //Journal of the Medical Library Association, 2004, vol.92, no.4, pp.429–437.
55. Rice S. Using big data to prevent drug errors. 2015 www.modernhealthcare.com/article/20150711/MAGAZINE/307119976
56. Lenfant C. Clinical Research to Clinical Practice — Lost in Translation? The New England Journal of Medicine, 2003, vol.349, pp.868–874, PMID: 12944573
57. Рост через инновации. Российский и международный опыт. www.pwc.ru/innovationsurvey
58. Big data in the healthcare industry: Growing Need for Computerized Decision Support. www.healthcare.siemens.com/magazine/mso-big-data-and-healthcare-1.html

59. Data-driven healthcare organizations use big data analytics for big gains. IBM Software White Paper. 2013. www03.ibm.com/industries/ca/en/healthcare/documents/Data.pdf
60. Davenport T.H., Barth P., Bean R. How 'Big Data' Is Different / MIT Sloan Management Review/ <http://sloanreview.mit.edu/article/how-big-data-is-different/>
61. Uijen A.A., Vande Lisdonk E.H. Multimorbidity in primary care: prevalence and trend over the last 20 years // *European Journal of General Practice*, 2008, vol.14, no.1, pp.28–32.
62. Лечение пациентов с множественными хроническими заболеваниями. Вестник McKinsey. <http://vestnikmckinsey.ru/healthcare-and-pharmaceuticals/Lechenie-patsiyentov-s-mnozhestvom-khronicheskikh-zabolevaniy>
63. Bates D.W., 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. <http://content.healthaffairs.org/content/33/7/1123.full?ijkey=QhqcLkIaGb84M>
64. Big Data and Analytics. www.mill-all.com/wp-content/uploads/2015/08/Healthcare
65. www.medicare.gov/
66. Mani N., Narayanan R., Raghunath M. Big Data in the healthcare provider space. www.mu-sigma.com/analytics/thought_leadership/decision-sciences.html
67. Challenges and opportunités with Big Data. <http://cra.org/ccc/wp-content/uploads/sites/2/2015/05/bigdatawhitepaper.pdf>
68. Hersh W. The Health Information Technology Workforce // *Journal Applied Clinical Informatics (ACI)*, 2010, vol.1, no.2, pp.197–212.
69. A Policy Forum on the Use of Big Data in Health Care, 2014. <http://bipartisanpolicy.org/library/policy-forum-use-big-data-health-care>
70. Robertson J., Dehart D., Tolle K., Hecherman D. Healthcare Delivery in Developing Countries: Challenges and Potential Solutions The fourth paradigm. 2009, pp.65–73
71. ABI Research: Integrating Consumer Wearable Health Devices Will Drive Healthcare Big Data Adoption. www.abiresearch.com/press/integrating-consumer-wearable-health
72. Мамедова М.Г. Информационная безопасность персональных медицинских данных в электронной среде // *Проблемы информационных технологий*, 2015, №2, pp.16–30. <http://jpit.az/index.php?mod=9&view=art&id=142>
73. Big data security and privacy issues in healthcare www.computer.org/csdl/proceedings/bigdatacongress/2014/5057/00/5057a762.pdf
74. Мəммədova M., Сəbrayılova Z. Böyük verilənlərin insan resurslarının idarə olunması məsələlərinin həllində tətbiqi, imkanları, problemləri // *İnformasiya texnologiyaları problemləri*, 2016, №1, s.39–48.
75. Mattison on Big Data: Great potential but risks exist. www.clinical-innovation.com/topics/analytics-quality/mattison-big-data-great-potential-risks-exist
76. Borkar VR, Carey MJ, Chen L: Big data platforms: what's next? // *ACM Crossroads* 2012, vol.19(no.1), pp.44–49.
77. Zikopoulos PC, DeRoos D, Parasuraman K, Deutsch T, Corrigan D, Giles J: *Harness the Power of Big Data*. The IBM Big Data Platform; McGraw-Hill: 2013, 281 p.
78. Alıquliyev R. M., İmamverdiyev Y. N., Abdullayeva F.C. Neft-qaz sənayesi üçün big data analitikanın cloud computing platformasında analytics-as-a-service kimi reallaşdırılma imkanlarının tədqiqi // *İnformasiya texnologiyaları problemləri*, 2016, №1, s.11–26.
79. Иэн Гортон, Джон Клейн. Системы для Больших Данных: конвергенция архитектур // «Открытые системы», 2015, №03. www.osp.ru/os/2015/03/13046898/
80. Big Data in the Cloud: Converging Technologies: How to Create Competitive Advantage Using Cloud-Based Big Data Analytics. / Intel IT Center. www.intel.com/content/dam/www/public/us/en/documents/product-briefs/big-data.pdf
81. Global Big Data spending in healthcare 2015 analysis and forecasts to 2019. www.prnewswire.com/news-releases/global-big-data-spending-in-healthcare-industry

82. Big Data&Analytics. IDC Future Scape: Worldwide Big Data and Analytics 2016 Predictions – APEJ Implications. Dec 2015 www.idc.com/prodserv/4Pillars/bigdata
83. The future of healthcare in Europe. The Economist Intelligence Unit Limited 2011, 54 p.
84. The Future of Healthcare Innovation, Big Data, Security and Patient Engagement. <http://electronichealthreporter.com/the-future-of-healthcare-innovation-big-data-security>
85. How to transform big data into better health. Workshop report, Italy, 2014. www.scienceeurope.org/uploads/publicdocumentsandspeeches/scspublicdocs/report
86. Continuous healthcare: Big data and the future of medicine. <http://venturebeat.com/2015/06/21/continuous-healthcare-bigdata-and-future-of-medicine/>

UOT 004. 62

Məmmədova Məsumə H.

AMEA İnformasiya Texnologiyaları İnstitutu, Bakı, Azərbaycan

masuma.huseyn@iit.ab.az

Elektron tibbdə Big Data: imkanlar, çağırışlar və perspektivlər

Tibbdə informasiyanın ifrat artımını təyin edən faktorlar nəzərdən keçirilmişdir. Tibbi verilənlərin spesifik xüsusiyyətləri tədqiq edilmiş, elektron tibbdə Big Data fenomeninin mahiyyəti və onun potensialı müəyyən edilmişdir. Big Data tətbiqlərin sistemləşdirilməsi aparılmış, müalicə-diaqnostik və idarəetmə qərarlarının dəstəklənməsində onların imkanları göstərilmişdir. Tibbin inkişafı və transformasiyası kontekstində bu sahədə BigData-analitikanın üstünlükləri qeyd edilmişdir. Big Data-nın tibbdə istifadəsini məhdudlaşdıran çağırışlar göstərilmiş və onların tədqiq olunan sahəyə tətbiqinin perspektiv istiqamətləri qeyd edilmişdir.

Açar sözlər: Big Data, elektron tibb, tibbi verilənlər, Big Data tətbiqlər, qərarların qəbulu, Big Data-çağırışlar.

Masuma H. Mammadova

Institute of Information Technology of ANAS, Baku, Azerbaijan

masuma.huseyn@iit.ab.az

Big Data in electronic medicine: opportunities, challenges and perspectives

The paper reviews the factors defining excessive growth of the information in medicine. The specific peculiarities of medical data are investigated; the essence of Big Data phenomenon and its potential in electronic medicine are defined. Big Data applications are systematized, and their capacity for the support of medical and diagnostic and management decisions is shown. Advantages of Big Data-analytics in this sphere are indicated in the context of development and transformation of medicine. The challenges, limiting the usage of Big Data in medicine, and the standpoint directions of its application in the given subject field are specified.

Keywords: Big Data, e-medicine, health data, Big Data-applications, decision-making, Big Data-challenges.