

Искусственный интеллект в оценке рисков развития сердечно-сосудистых заболеваний

А.В. Гусев¹, Т.Ю. Кузнецова², И.Н. Корсаков¹

¹ ООО «Комплексные медицинские информационные системы (К-МИС)»

² ФГБОУ ВО «Петрозаводский государственный университет», Петрозаводск, Российская Федерация

Для корреспонденции:

agusev@kmis.ru

Artificial intelligence for cardiovascular risks assessment

A.V. Gusev¹, T.Yu. Kuznetsova², I.N. Korsakov¹

¹ Complex Medical Information Systems Company (K-MIS)

² Petrozavodsk State University, Petrozavodsk, Russian Federation

Cardiovascular disease (CVD) is one of the leading causes of death and disability in most countries of the world, including Russia. According to Russian Federal State Statistics Service (Rosstat) in 2016, 904 thousand people died from diseases of the circulatory system in Russia, which amounted to 47.8% in the structure of the causes of mortality. A significant proportion of this morbidity and mortality could be prevented through early diagnosed and prevention strategies, both for people with established disease and for those at high risk of developing disease. The challenge is to ensure a smooth transition from inpatient to outpatient care and to improve the education of all patients. Despite the current world clinical guidelines, which clearly indicate the sequence of actions of the doctor, including the obligation to objectively assess the data on the state of health, to identify risk factors, to determine the cardiovascular risk in a particular patient, and then offer treatment to reduce this risk, but the mass full-fledged identification of risk factors and the overall risk assessment of CVD is not currently carried out in Russia. Artificial Intelligence (AI) tools have advanced enough so that they can be integrated into decision support systems for real applications and are impacting decision making in substantive ways. This paper reviews decision making theories and AI tools and the intelligent decision systems that result from the integration of these concepts. This system could be embedding into any medical information system with electronic medical records (EHR) and using data mining, data extraction including data from unstructured medical records, to provide automatic identification of risk factors and risk assessment of cardiovascular diseases of the patient.

Key words: Electronic health record (EHR), cardiovascular diseases (CVD), risk factors, clinical decision support system (CDSS), decision support system (DSS), medical decision-support systems (MDSS), artificial intelligence (AI)

Сердечно-сосудистые заболевания (ССЗ) являются одной из главных причин смертности и инвалидизации в большинстве стран мира [1], в т. ч. в России. По данным Росстата в 2016 г. от болезней системы кровообращения в России умерло 904 тыс. человек, что составило 47,8 % в структуре причин смертности населения. Основными причинами смерти явились ишемическая болезнь сердца (ИБС, количе-

ство умерших 481 тыс. человек, из них вследствие инфаркта миокарда 62,9 тыс. чел.) и цереброваскулярные болезни (279,8 тыс. человек) [2].

Ключевая причина таких высоких показателей – это низкая выявляемость болезней на ранних стадиях, недостаточная приверженность пациентов высокого риска рекомендациям врачей по профилактике.

В большинстве развитых стран, где длительное время уделяли большое внимание совершенствованию ►

мер предотвращения заболеваний, а не только их лечению, произошли значительные положительные сдвиги в показателях смертности от ССЗ, что в результате и определило большой отрыв в продолжительности жизни и эффективности борьбы с основными хроническими неинфекционными заболеваниями.

Мировая индустрия охраны здоровья проходит сейчас через поистине революционные преобразования. Меняются не только технологии, методики и стандарты лечения, но и сам подход к организации здравоохранения. В частности, в странах-лидерах отчетливо наблюдается смена парадигмы в сторону активной профилактики и предотвращения, а не лечения, заболеваний и стремление к сокращению стационарной помощи в пользу амбулаторного лечения, ухода на дому и самостоятельной заботы пациентов о собственном здоровье.

Например, в системе здравоохранения Великобритании давно определили вторичную профилактику как ключевую возможность предотвратить возникновение дорогостоящих осложнений. Для практикующих врачей первичной медицинской помощи предусмотрены материальные стимулы для оказания целого ряда превентивных мероприятий, включая оценку рисков развития хронических заболеваний. Например, ведение пациентов с диабетом включает регулярные оценки индекса массы тела, уровня HbA1c и обследования стопы. Стандартизация клинических оценок с помощью контрольных чек-апов гарантирует, что пациенты с риском развития осложнений выявляются на ранней стадии, что, в свою очередь, позволяет получить ответ на лечение, основанный на потребностях, и в конечном итоге улучшить качество обслуживания пациентов [3].

В большинстве современных мировых клинических рекомендациях четко указана последовательность действий врача, к которому за помощью обратился пациент, включая обязанность оценить объективные данные здоровья, выявить факторы риска и на основании их определить сердечно-сосудистый риск у конкретного пациента, а затем предпринимать шаги по снижению этого риска [4].

У выполнения этих рекомендаций в России есть несколько существенных барьеров, которые, по сути, сводят на нет ее массовое соблюдение практическим здравоохранением:

1. На сегодня только по оценке сердечно-сосудистого риска известно более 40 различных шкал, рекомендаций, методик. Запомнить и удер-

жать в памяти все эти шкалы невозможно.

2. Типичный российский врач — это очень усталый человек, который много перерабатывает. Согласно всероссийскому исследованию Академии труда и социальных отношений, проведенному в 2016–2017 годах, около 41% российского медперсонала работают более 60 часов в неделю. Почти 70% врачей отмечают увеличение объема нагрузки без увеличения численности персонала, а 50% опрошенных врачей жалуются на увеличение объема работы, не связанной с лечением пациентов [5].

3. На рутинном приеме просто нет времени на полноценную оценку рисков. Врачам приходится тратить большое количество времени на оформление документации и отчетов, включая соблюдения требований, продиктованных необходимостью формировать реестры по ОМС и собирать данные для сдачи большого количества медицинской статистики. Показано, что врачам приходится посвящать бумажной работе основную часть рабочего времени: из 12 минут приема 70–80% времени может занять только ввод данных [5].

4. На практике доктор не всегда информирует пациента о вариантах и рисках лечения, доступном выборе и потенциальной отдаче, так как это затягивает прием и приводит к более критическому восприятию информации пациентом. Отсутствует эффективная система мотивации врачей к оценке риска, фактически это делают лишь единичные энтузиасты, и то только из числа имеющих достаточно свободного для этого времени.

В итоге массового полноценного выявления факторов риска и общей оценки риска развития ССЗ не проводится. Простые призывы к врачебной ответственности или нормативное закрепление этих требований вряд ли на практике существенно изменят сложившуюся ситуацию, т.к. они не меняют причины этой проблемы.

■ АНАЛИЗ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ ПРОГНОЗИРОВАНИЯ ЗАБОЛЕВАНИЙ И ОЦЕНКИ ФАКТОРОВ РИСКА

Интеллектуальные системы, построенные на базе машинного обучения и технологий искусственного интеллекта, продемонстрировали большую перспективность в прогнозировании и выявлении угроз общественному здравоохранению, а также улучшение результатов ведения пациентов высокого риска. По мере того, как они

продолжают совершенствоваться, медицинские работники будут все больше и больше использовать этот мощный инструмент для оказания пациентам более точной, своевременной профилактической помощи [7, 6]. По оценкам Института Слоана Кеттеринга, изучившего эффективность диагностики и лечения онкологических больных, врачи используют только 20% доступной информации [8]. Применяя развитые алгоритмы, способные обрабатывать огромные объемы данных и в течение нескольких секунд предоставлять врачу всестороннюю оценку имеющейся медицинской информации, можно существенно повысить эффективность работы врача, при этом не удлиняя время приема и даже сокращая нагрузку.

Имеются следующие предпосылки для такого подхода:

1. Во всем мире и в России в частности идет массовый переход на ведение электронных медицинских карт (ЭМК). Уже сейчас подавляющее большинство рабочих мест врачей в России оснащено ЭМК различных производителей. Согласно государственной программы «Создания единого цифрового контура в здравоохранении на основе ЕГИСЗ», к 2020 г. 80% врачей будут вести свою работу в юридически значимом электронном медицинском документообороте.

2. Исходные данные, необходимые для определения риска хотя бы по известным методикам наподобие шкалы Score, имеются в подавляющем большинстве применяемых в России систем ведения электронной медицинской карты пациента.

Таким образом, одним из самых перспективных способов повышения эффективности профилактических методов снижения заболеваемости и в итоге смертности, является создание систем поддержки принятия врачебных решений (СППВР), которые можно было бы встраивать в медицинские информационные системы с системой ведения электронных медицинских карт (ЭМК). Такие СППВР, построенные как централизованные сервисы по анализу ЭМК и выявлению в них факторов риска или подозрений на наличие заболеваний на ранней стадии, могли бы взять часть профилактической работы на свои плечи, причем в полностью автоматическом режиме.

Барьерами на пути создания и массового внедрения соответствующих программных продуктов являются:

1. Действующее в России законодательство, согласно которому программное обеспечение, предназначенное разработчиком для принятия врачами решений и/или прямо влияющее на такое

решение, относится к медицинским изделиям и подлежит непростой, затратной и длительной процедуре государственной регистрации, включающей технические и клинические испытания [9].

2. Длительность, сложность и затратность разработки таких систем, т.к. для этого требуется привлечение опытных клиницистов и медицинских экспертов, причем не только на этапе постановки задачи – но и на этапе тестирования и оценки результатов опытных эксплуатаций. Такие специалисты редки и заняты.

В этой связи целесообразным является создание систем поддержки принятия врачебных решений в виде самостоятельного программного продукта, построенного по принципу открытого web-сервиса, который можно было бы использовать в любых системах ведения электронных медицинских карт путем интеграции с API.

Такая информационная система должна принимать для анализа данные из электронной медицинской карты и самостоятельно анализировать эти данные о состоянии здоровья пациента и, применяя различные методики анализа, включая существующие шкалы оценки риска и методы искусственного интеллекта, выявить персональные факторы риска пациента и затем передать их в формализованном виде в ту информационную систему, которая вызвала этот сервис.

Такой подход позволит обеспечить максимальную скорость в обновлении программного продукта и доставки улучшенной функциональности до каждого рабочего места врача. В частности, он позволит ускоренными темпами внедрять в практическое звено не только известные, но и новые методы оценки рисков, в том числе построенные с использованием методов машинного обучения на больших объемах накопленных данных. За счет применения искусственного интеллекта такой сервис научится со временем обнаруживать в медицинских данных новые взаимосвязности, которые можно использовать для превентивных вмешательств и более выверенных мер профилактики, включая более эффективное использование новейшей диагностики и методов лечения [10].

Еще одно важное преимущество создания информационной системы в качестве независимого открытого web-сервиса – это его способность осуществлять оценку данных не дожидаясь осмотра пациента врачом и не отвлекая врача. Вызов такого сервиса могли бы осуществлять централизованные информационные системы в сфере здравоохранения по различным событиям: например при поступлении новых результатов ►►

лабораторного исследования, периодическое обновление рисков по мере обучения сервиса или по расписанию. Это позволило бы постоянно актуализировать индивидуальные и популяционные оценки рисков, не дожидаясь обращения больных за медицинской помощью в учреждения здравоохранения. Многие современные медицинские информационные системы включают в себя Личный кабинет Пациента, в который данные информация может быть доставлена, а в экстренных случаях может быть послано дополнительное оповещение.

Несмотря на описанные преимущества, создание эффективного сервиса для оценки рисков развития заболеваний пациента на основе анализа электронной медицинской карты имеет ряд серьезных проблем. Главная из них связана с текущим неудовлетворительным уровнем развития удобства ведения ЭМК. Опрос более чем 1100 специалистов, проведенный в 2017 г. в США, показал, что 61% респондентов считают, что ЭМК вызвали очень плохую отдачу от цифровых инвестиций, и только 10% оценили рентабельность инвестиций (ROI) от ЭМК как положительную. Громоздкие рабочие процессы, неудобные и недружественные интерфейсы, перегруженные управляющими элементами, неинтуитивные правила работы являются привычными недостатками очень многих ЭМК во всем мире, в том числе и созданные лидерами рынка. Такие системы нередко добавляют к рабочим нагрузкам врачей дополнительное бремя и значительно сокращают их время взаимодействия с пациентами. Встраивание в ЭМК эффективных сервисов поддержки принятия врачебных решений могло бы частично оправдать сложности ведения ЭМК.

Проведенный в США в 2017 г. опрос 4197 практикующих врачей, показал, что 26% из 1792 врачей испытывали выгорание от перехода на ЭМК. 70% медицинских работников, которые использовали электронные медицинские карты, испытывали стресс, вызванный этой необходимостью. Наиболее высокие показатели наблюдались среди специалистов первичного звена. Риск выгорания из-за применения ЭМК был повышен среди врачей, которые отмечали нехватку времени для ведения медицинской документации, по сравнению с теми, кто имел достаточное количество времени, а также среди тех, кому приходи-

лось тратить много времени дома на работу с электронными медицинскими картами и тех, кто согласен с тем, что электронные медицинские карты способствуют ежедневной фрустрации [12].

Еще одна проблема связана со способом хранения электронных медицинских записей. Достаточно часто разработчики, стремясь снизить нагрузку на врачей по их ведению и, как следствие, сопротивляемость медицинского персонала, допускают достаточно вольные способы внесения записей, в том числе повсеместно используются текстовые шаблоны, копирование и вставка текста, повторные автоматические цитирования уже ранее внесенной информации. Несмотря на развитие стандартов медицинской информатики, классификаторов и онтологий для машинно-понимаемого хранения данных, надо признать – что они не получили массового признания и применения в системах ЭМК. Поэтому большинство из накапливаемых медицинских записей – это самые обычные неструктурированные (неформализованные) текстовые поля, содержащие произвольно написанный текст с сокращениями, орфографическими и пунктуационными ошибками, жаргонами и т.д. Такой текст без специальной предварительной обработки непригоден для эффективного анализа.

Несмотря на это, различные исследователи и ИТ-гиганты работают над изучением и совершенствованием обработки медицинских данных из неструктурированных электронных медицинских карт и их успехи вселяют надежду, что в целом именно такой подход обеспечит в будущем наилучшее сочетание эффективности и простоты, так необходимых для массового внедрения.

Отличительной особенностью предварительной обработки неструктурированной медицинской информации является тот факт, что вся имеющаяся информация использует синтаксис, семантику и грамматику Русского языка. На сегодняшний момент нет даже медицинского корпуса Русского языка. Очевидно, что для решения этих задач требуются системные исследовательские работы, которые могут быть выполнены на основе запроса разработчиков программного обеспечения.

В целом архитектура сервиса, осуществляющего оценку ЭМК и формирование заключения о рисках развития заболеваний пациента, выглядит так, как показано на рисунке.

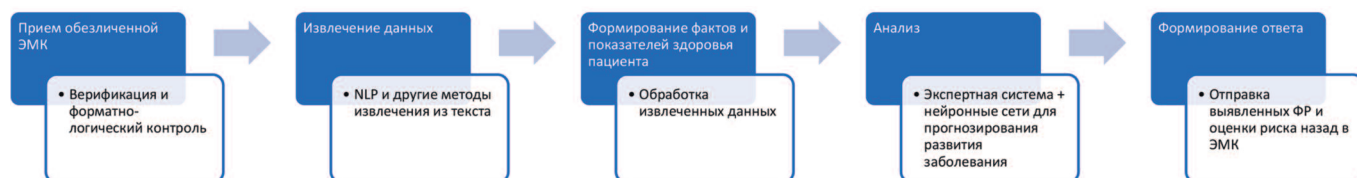


Рис. Последовательность обработки информации в системе поддержки принятия врачебных решений, предназначенной для прогнозирования заболеваний и оценки рисков

Анализ выявленных показателей здоровья и фактов может осуществляться 2-мя основными способами:

1. На основе экспертной системы. Это специальный блок, созданный обычными методами программирования на основании алгоритмов, описанных в уже проведенных исследованиях и методиках оценки рисков – например шкалы Score, исследования Фрамингем и т.д.

2. На основе машинного обучения. Это дополнительный блок, который осуществляет оценку рисков на основании проведенного ранее машинного обучения, выполненного на больших объемах предварительно собранных и размеченных электронных медицинских карт пациентов схожего поло-возрастного состава, этнической и территориальной принадлежности.

Применение одновременно двух подходов является принципиально важным, потому что они имеют разную природу и ценность. Построенные на основании экспертного подхода будут иметь максимальную репрезентативность и понятность для лечащего врача, т.к. оценки на их основе будут понятны клиницисту, их можно снабдить ссылками на публикации в Интернет. Однако они имеют минимальную персональность и точность, т.к., как правило, построены на очень ограниченном наборе входных данных (возраст, АД, липидный спектр и другие, хорошо изученные факторы риска) и не учитывают множество личных особенностей пациента.

Оценки, построенные с помощью машинного обучения, имеют обратные минусы и плюсы. Как правило созданные с использованием нейронных сетей, алгоритмы представляют из себя «черный ящик» и не могут объяснить причинно-следственные связи врачу – почему система дала ту или иную оценку. Вместе с этим такие алгоритмы могут дать оценку более точную, чем построенные на базе обычной регрессионной математике [11].

Для прогнозирования конкретного события (ухудшение или манифест заболевания, госпитализация или смерть), а не научного исследования связи между конкретными факторами риска и клинически интерпретируемым событием, оценки данных пациента на основе методик, построенных по результатам исследований на основе традиционных регрессионных моделей, является в данное время недостаточным. Они должны быть дополнены моделями, которые обеспечивают более гибкую связь между переменными-показателями и результатом. Эти модели, построенные на основе методов Machine learning (ML) (машинного обучения) и Deep learn-

ing (DL) (глубокого обучения), каждый из которых зарекомендовал себя для одного или другого применения. Результаты исследований свидетельствуют, что они позволяют получить лучшее окончательное прогнозирование, хотя и жертвуя интерпретацией факторов риска связанных с результатом [11].

При этом важно отметить, что нет единого, общепринятого и эффективного подхода к выбору различных методов машинного обучения. В настоящее время развитие ИИ, возможностей аппаратного обучения и математических алгоритмов достигло такого уровня, что существует очень много различных вариантов решения задачи прогнозирования события со здоровьем пациента. Каждый из них может быть эффективен или мало полезен в зависимости от конкретной задачи, поэтому перед разработчиком стоит серьезная задача изучить наиболее известные из них и проверяя каждый метод, найти лучшее решение для своей конкретной задачи.

■ ЗАКЛЮЧЕНИЕ

Обобщая все вышесказанное, мы пришли к следующим выводам:

1. Для снижения заболеваемости и смертности от сердечно-сосудистых заболеваний наиболее эффективным является массовое внедрение методов профилактики в первичное звено здравоохранения. Для этого необходимо создание системы поддержки принятия врачебных решений (СППВР), способной выявлять факторы риска и подозрения на наличие заболеваний на как можно более ранней стадии, анализируя данные электронной медицинской документации.

2. СППВР должна быть встроена в систему ведения электронных медицинских карт и уметь работать в том числе с неструктурированными медицинскими записями, «извлекая» из нее данные, необходимые для оценки факторов риска.

3. В работе СППВР должны применяться, как оценка данных на основе экспертных систем, так и методы искусственного интеллекта, в частности машинное обучение – для обеспечения сбалансированной достоверности, точности и репрезентативности тех результатов, что СППВР будет выводить врачу.

Исследование не имело спонсорской поддержки.

Авторы заявляют об отсутствии конфликта интересов. █