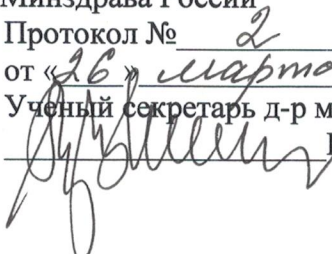


**ФГБУ «ЦЕНТРАЛЬНЫЙ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ИНСТИТУТ  
ОРГАНИЗАЦИИ И ИНФОРМАТИЗАЦИИ ЗДРАВООХРАНЕНИЯ»  
МИНИСТЕРСТВА ЗДРАВООХРАНЕНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ**

---

**ПРИНЯТО**

Ученым советом ФГБУ «ЦНИИОИЗ»  
Минздрава России  
Протокол № 2  
от «26» марта 2024 г.  
Ученый секретарь д-р мед. наук, доцент  
  
В.В. Люцко

**УТВЕРЖДАЮ**

Директор ФГБУ «ЦНИИОИЗ» Минздрава  
России, д-р мед. наук, профессор

О.С. Кобякова

«26» марта 2024 г.



**ЭТИЧЕСКАЯ ЭКСПЕРТИЗА  
КЛИНИЧЕСКИХ ИСПЫТАНИЙ  
СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА**

Методические рекомендации № 03-24

г. Москва - 2024

УДК

ББК

**Авторы:**

Гусев А.В., к.т.н., старший научный сотрудник отдела научных основ организации здравоохранения ФГБУ «ЦНИИОИЗ» Минздрава России, директор по развитию К-Скай

Кобякова О.С., д.м.н., профессор, директор ФГБУ «ЦНИИОИЗ» Минздрава России

Ковширина Ю.В., к.м.н., ведущий научный сотрудник отдела научных основ организации здравоохранения ФГБУ «ЦНИИОИЗ» Минздрава России

Реброва О.Ю., д.м.н., профессор кафедры медицинской кибернетики и информатики им. С.А. Гаспаряна медико-биологического факультета ФГАОУ ВО «РНИМУ им. Н.И. Пирогова» Минздрава России

**Рецензенты:**

Хохлов А.Л., д.м.н., профессор, академик РАН, ректор ФГБОУ ВО ЯГМУ Минздрава России

Зарубина Т.В., д.м.н., профессор, член-корреспондент РАН, Заслуженный деятель науки Российской Федерации, директор института цифровой трансформации медицины, заведующая кафедрой медицинской кибернетики и информатики медико-биологического факультета ФГАОУ ВО «РНИМУ им. Н. И. Пирогова» Минздрава России

## ТЕРМИНЫ И СОКРАЩЕНИЯ

Термин / Сокращение	Определение / Наименование
БД	База данных
ГОСТ	Государственный стандарт
ЕГИСЗ	Единая государственная информационная система в сфере здравоохранения
ИС	Информационная система
ИИ	Искусственный интеллект
КИ	Клиническое испытание
МИ	Медицинское изделие
Минздрав России	Министерство здравоохранения Российской Федерации
МИС МО	Медицинская информационная система медицинской организации
ММО	Модель машинного обучения
МО	Медицинская организация
ПО	Программное обеспечение
СИИ	Система искусственного интеллекта
СППВР	Система поддержки принятия врачебных решений

## СОДЕРЖАНИЕ

<b><u>ВВЕДЕНИЕ .....</u></b>	<b><u>5</u></b>
<b><u>1. ОБЩИЕ СВЕДЕНИЯ О СИСТЕМАХ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ ЗДРАВООХРАНЕНИЯ.....</u></b>	<b><u>6</u></b>
<b><u>2. ОБЗОР ОСНОВНЫХ ЭТИЧЕСКИХ ПРОБЛЕМ, ВОЗНИКАЮЩИХ ПРИ ИСПОЛЬЗОВАНИИ СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА.....</u></b>	<b><u>8</u></b>
2.1. НЕДОСТАТОЧНОСТЬ ДОКАЗАТЕЛЬСТВ ЭФФЕКТИВНОСТИ И БЕЗОПАСНОСТИ.....	8
2.2. ВЫСОКИЙ РИСК ДЕГРАДАЦИИ МЕТРИК ТОЧНОСТИ СИИ ПРИ ЕЕ ПРИМЕНЕНИИ В РЕАЛЬНОЙ КЛИНИЧЕСКОЙ ПРАКТИКЕ.....	11
2.3. ПРОБЛЕМА ОТВЕТСТВЕННОСТИ ЗА ОШИБКИ В РАБОТЕ СИИ.....	12
2.4. ПРОБЛЕМА «ЧЕРНОГО ЯЩИКА».....	12
2.5. СТРАХИ ПЕРЕД СИИ.....	12
2.6. КОНФИДЕНЦИАЛЬНОСТЬ МЕДИЦИНСКОЙ ИНФОРМАЦИИ .....	13
<b><u>3. РЕКОМЕНДАЦИИ ПО ЭТИЧЕСКОЙ ЭКСПЕРТИЗЕ КЛИНИЧЕСКИХ ИСПЫТАНИЙ СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА.....</u></b>	<b><u>13</u></b>
3.1. ОЦЕНКА ОБОСНОВАННОСТИ РАЗРАБОТКИ .....	15
3.2. ОЦЕНКА ДИЗАЙНА И МЕТОДОЛОГИЧЕСКОГО КАЧЕСТВА КЛИНИЧЕСКОГО ИСПЫТАНИЯ.....	16
3.3. ОЦЕНКА УРОВНЯ ЗРЕЛОСТИ КОМПАНИИ-РАЗРАБОТЧИКА.....	17
3.4. ОЦЕНКА ДОВЕРИЯ К НАБОРАМ ДАННЫХ, ИСПОЛЬЗОВАННЫМ ДЛЯ МАШИННОГО ОБУЧЕНИЯ	18
3.4.1. ОЦЕНКА ИСТОЧНИКОВ, ИСПОЛЬЗОВАННЫХ ДЛЯ СОЗДАНИЯ НАБОРОВ ДАННЫХ .....	18
3.4.2. ОЦЕНКА НАБОРОВ ДАННЫХ.....	19
3.5. ОЦЕНКА ДОВЕРИЯ К МОДЕЛЯМ МАШИННОГО ОБУЧЕНИЯ.....	25
<b><u>ЗАКЛЮЧЕНИЕ .....</u></b>	<b><u>28</u></b>
<b><u>СПИСОК ЛИТЕРАТУРЫ .....</u></b>	<b><u>29</u></b>

## **ВВЕДЕНИЕ**

Здравоохранение является одной из приоритетных отраслей для практического применения систем искусственного интеллекта (ИИ), что определено «Национальной стратегией развития искусственного интеллекта на период до 2030 года».

Внедрение в медицинских организациях (МО) систем искусственного интеллекта (СИИ) может способствовать повышению уровня жизни населения и качества медицинской помощи, включая профилактику, диагностику и прогнозирование возникновения и развития заболеваний, подбор лекарственной терапии, сокращение угроз пандемий, автоматизацию и т.д. [1].

С точки зрения российского законодательства СИИ, предназначенная производителем для применения при оказании медицинской помощи медицинскими работниками, является медицинским изделием (МИ). Для вывода такой СИИ на рынок компания-разработчик должна пройти процедуру государственной регистрации в установленном законодателем порядке. Важнейшим этапом этого процесса является проведение клинических испытаний (КИ), предназначенных для подтверждения эффективности и безопасности СИИ.

В последнее время в профессиональном сообществе сформировалось общее понимание, что перед проведением КИ СИИ обязательным мероприятием является получение одобрения этического комитета на предложенную программу КИ. Технологии ИИ обладают рядом ключевых особенностей, которые существенным образом влияют на проведение этической экспертизы. Сотрудникам этических комитетов, рассматривающих программу КИ на СИИ, необходимо хорошо разбираться как в нюансах технологий ИИ, включая процессы подготовки данных и машинного обучения, так и в особенностях применения СИИ в реальной клинической практике. Без знания таких особенностей проведение этической экспертизы может быть осуществлено неправильно или некачественно.

Целью данных методических рекомендаций является предоставление сотрудникам этических комитетов необходимых знаний об особенностях СИИ с точки зрения биоэтики.

## 1. ОБЩИЕ СВЕДЕНИЯ О СИСТЕМАХ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ ЗДРАВООХРАНЕНИЯ

Внедрение решений на основе технологий ИИ являются одним из основных направлений цифровой трансформации здравоохранения. Традиционно создание СИИ возможно на основе следующих подходов:

- на основе инженерии знаний. В этом случае для анализа и интерпретации медицинских данных используются алгоритмы, созданные разработчиками и экспертами на основании знаний по предметной области из литературных источников.
- на основе данных. В этом случае для анализа и интерпретации медицинских данных используются модели, полученные методами машинного обучения на основании обучающих наборов данных.
- на основе гибридного подхода. В этом случае при создании СИИ используются и методы инженерии знаний и методы машинного обучения.

Созданное с помощью технологий ИИ программное обеспечение (ПО) может быть ориентировано на решение различных клинических задач, включая поддержку принятия врачебных решений (СППВР) при диагностике, лечении, реабилитации и профилактике.

Считается, что внедрение СИИ произведет революцию в здравоохранении, в т.ч. благодаря консультативной помощи при принятии клинических решений, анализе цифровых изображений, управлении рисками, прогнозировании и контроле безопасности назначаемого лечения [2].

Важнейшим глобальным драйвером ИИ для здравоохранения являются существенные инвестиции, вкладываемые во всем мире в создание и развитие различных цифровых продуктов и сервисов. По данным CB Insights [4], в 2021 г. объём суммарных инвестиций в компании, предлагающие здравоохранению решения на основе ИИ, составил 12,2 млрд долл. Для сравнения, в 2020 г. эта цифра равнялась 6,627 млрд долл., в 2019 г. — 4,129 млрд долл., а в 2018 г. — 2,7 млрд долл. Прогнозируется, что к 2026 г. они увеличатся до 49,10 млрд долл. при среднегодовом темпе роста 23,46% [3].

Вторым по значимости драйвером являются научные исследования. Первые публикации по теме применения технологий ИИ для решения задач здравоохранения появились в рецензируемой научной медицинской литературе в 50-х годах прошлого века. Долгое время динамика исследований была довольно медленной, однако, начиная с 2017 г., наблюдается существенный рост публикаций, в которых изучались или применялись различные технологии ИИ, включая машинное обучение (рис. 1).

Третьим драйвером является развитие непосредственно самих информационных технологий (ИТ): рост вычислительной мощности компьютеров, развитие облачных вычислений, быстрое накопление больших данных. Это служит основой для применения технологий машинного обучения, включая компьютерное зрение, работу с естественным языком, прогнозную аналитику и т.д.

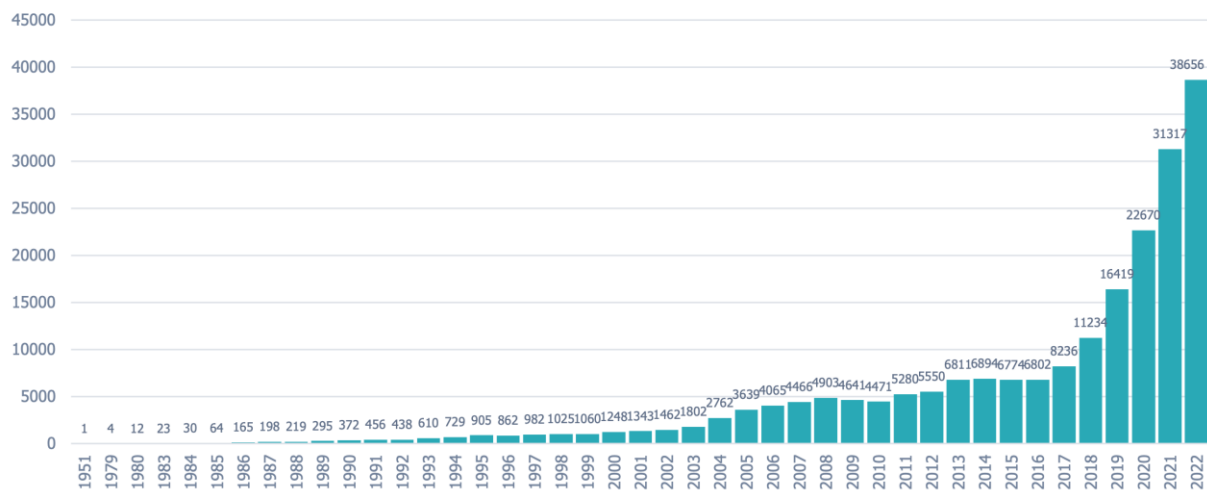


Рисунок. 1. Динамика рецензируемых научных публикаций, в которых изучались или использовались технологии искусственного интеллекта в 1951-2022 гг. (по данным PubMed)

Проникновение СИИ затрагивает в настоящее время практически все направления работы медицинских организаций и органов управления и надзора в сфере здравоохранения, а также все основные нозологические группы и этапы лечебно-диагностического процесса. Лидирующими из них являются медицинская диагностика и анализ изображений, прогнозная клиническая аналитика, системы поддержки принятия решений, анализ данных носимых устройств, телемедицина, мониторинг пациентов с хроническими неинфекционными заболеваниями, виртуальные ассистенты, включая робот-ассистированную хирургию, помощь в решении проблем психического здоровья и т.д. [5, 6, 7].

В последнее время значительный прогресс при создании и внедрении СИИ наблюдается в интерпретации медицинских данных, включая медицинские изображения, с использованием искусственных нейронных сетей и методов глубокого обучения [8, 9].

Основной предпосылкой интереса к широкому внедрению СИИ в практическом здравоохранении является рост заболеваемости, в особенности хроническими

неинфекционными заболеваниями. Вызванный этим рост затрат на здравоохранение создает необходимость в более эффективном отраслевом управлении и использовании ограниченных ресурсов, в т.ч. благодаря автоматизации рутинных функций, сокращению врачебных ошибок, раннему выявлению опасных заболеваний и последующему более эффективному лечению пациентов [10].

Применение технологий ИИ для улучшения различных аспектов клинической практики, получившее в зарубежной литературе также название дополненной медицины (augmented medicine), предполагает конвергенцию человеческого и искусственного интеллекта [11, 12, 13].

Одновременно в последнее время в профессиональном сообществе и научной литературе мы видим рост обсуждений и исследований, связанных с оценкой безопасности и эффективности, необходимых для широкого внедрения СИИ в практическое здравоохранение [14, 15].

## **2. ОБЗОР ОСНОВНЫХ ЭТИЧЕСКИХ ПРОБЛЕМ, ВОЗНИКАЮЩИХ ПРИ ИСПОЛЬЗОВАНИИ СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА**

Основными проблемами внедрения и использования СИИ в условиях практического здравоохранения являются [16, 17, 18, 19]:

- недостаток доказательств эффективности и безопасности этих решений с точки зрения допуска их для применения медицинскими работниками,
- повышенный риск причинения вреда здоровью пациентов в силу потенциальных ошибок обработки данных и формируемых заключений (рекомендаций),
- сложность интерпретации решений при машинном обучении (проблема «черного ящика»),
- повышенные риски самообучающихся алгоритмов, которые способны изменять свою работу вследствие появления новых клинических данных, в том числе получаемых во время их эксплуатации,
- проблемы кибербезопасности, включая несанкционированное вмешательство в алгоритмы работы СИИ или доступ к персональным данным пациентов,
- проблемы смещения данных, которые в свою очередь приводят к асимметрии между теми данными, на которых обучались ИИ-модели и данными, которые они анализируют при применении в условиях реальной клинической практики.

### **2.1. Недостаточность доказательств эффективности и безопасности**



Несмотря на проведение сотен контролируемых испытаний и десятков систематических обзоров по теме оценки клинической эффективности различных моделей машинного обучения (ММО), тем не менее имеется дефицит доказательств клинической эффективности и безопасности разрабатываемых и выводимых на рынок СИИ в отношении не только клинических, но и суррогатных исходов пациентов, а также потребления ресурсов здравоохранения [20]. Так, L. Wynants и соавт., 2020, представили результаты систематического обзора достоверности и полезности моделей прогнозирования развития COVID-19, в котором авторы сделали вывод, что многие изученные прогностические модели имели недостатки в части описания процесса их создания, содержали некоторый риск систематических ошибок, а также имели метрики точности, вероятнее всего, слишком оптимистичные [21].

К аналогичным выводам пришли M. Roberts и соавт., 2021, которые проанализировали результаты 62 исследований, посвященных разработке моделей машинного обучения для анализа рентгенограмм грудной клетки и изображений компьютерной томографии (КТ) грудной клетки для выявления признаков COVID-19, в котором авторы сделали вывод, что все изученные модели содержали ряд методологических недостатков и предубеждений [22].

A. Wong и соавт., 2021, опубликованы результаты независимого исследования точности работы модели на основе ИИ для прогнозирования сепсиса. Данная модель входит в число 20 запатентованных ИИ-алгоритмов, предназначенных для выявления и прогнозирования различных заболеваний, созданных компанией Epic Systems, крупнейшего американского поставщика одноименной системы ведения электронных медицинских карт. Исследование было проведено в условиях реальной клинической практике Мичиганского университета, где были изучены 38 тыс. историй болезни. Авторы пришли к выводу, что в реальности чувствительность модели составила 33%, точность 12%. Из 6 971 предупреждений системы о наличии сепсиса только 843 были правильными [23].

Таким образом, ряд исследований и систематических обзоров литературы показывают, что в настоящее время проблема полноценной оценки клинической эффективности и безопасности СИИ является весьма актуальной.

В России эта проблема также актуальна. Согласно действующему законодательству, СИИ подлежат государственной регистрации в качестве медицинского изделия [24], для чего, в свою очередь, необходимо проведение клинических испытаний (КИ). Как правило, программа клинических испытаний создается разработчиком СИИ совместно с внешней

медицинской организацией, имеющей в соответствии с действующим нормативным регулированием право на проведение таких испытаний. Целью КИ является оценка эффективности и безопасности СИИ. На практике КИ СИИ выполняется чаще всего в формате внешней валидации точности работы встроенных в СИИ моделей машинного обучения. Такой подход позволяет оценить, будут ли достигнуты заявленные производителем метрики точности работы модели на данных, которые не были использованы при создании или тестировании такой системы [25].

Любое программное обеспечение потенциально может содержать ошибки, связанные как с недостаточным качеством разработки и тестирования, так и с особенностями оборудования, поданными для анализа данными или применения его пользователями. Существуют убедительные доказательства, что применение ПО при оказании медицинской помощи может действительно приводить к причинению вреда, в том числе летальным исходам. Так, Guy Martin и соавт., 2019, приводят результаты ретроспективного анализа 10-летних данных об инцидентах в Англии и Уэльсе, связанных с применением ПО для медицинских целей. Авторы выявили 2627 инцидентов, связанных со сбоями в таком ПО. Из них 82% не причинили вреда пациентам, 13% причинили незначительный вред, 4% причинили вред средней степени тяжести, 1% причинили серьезный вред и менее 1% способствовал смерти больного. При этом 75% инцидентов были предотвратимыми. Авторы пришли к выводу, что ошибки в ПО несомненно являются потенциальным источником вреда, вследствие чего необходимо рассматривать ПО для здравоохранения с точки зрения фундаментальных принципов безопасности пациентов, разрабатывать более эффективные методы выявления последствий ошибок и принимать меры для снижения их вероятности и риска [26].

Отдельно следует упомянуть о безопасности таких систем в связи с судебным случаем компании “Practice Fusion” (США) [27]. Компания получала вознаграждение от фармацевтической компании «Purdue Pharma» в обмен на внедрение специальной всплывающей подсказки в разработанную программу для ведения электронных медицинских карт. Подсказка подводила врачей к клиническим действиям, результатом которых стало непропорционально высокое количество выписанных рецептов на опиоидный препарат пролонгированного действия ОхуContin. С 2016 по 2019 г., когда преступный сговор программистов и фармацевтов попал в поле зрения органов правопорядка, модуль поддержки клинического решения “Practice Fusion” всплывал на компьютерах лечащих врачей 230 миллионов раз [20].

Для создания подавляющего большинства современных СИИ разработчики используют методы и алгоритмы машинного обучения на основе специального подготовленных наборов данных, которые могут содержать ряд серьезных проблем с качеством, что приводит к повышению риска причинения вреда здоровью пациента по сравнению с обычным программным обеспечением, не использующем технологии ИИ.

Например, при создании СИИ нередко используются открытые международные наборы данных, в которых отсутствуют или слабо представлены записи о пациентах из российской популяции, что может приводить к низкой обобщаемости (внешней валидности) таких систем.

Зачастую разметка данных осуществляется сотрудниками компаний-разработчиков с неизвестной (возможно – недостаточной) компетенцией. В итоге наборы данных могут содержать систематические отклонения или явные ошибки, которые очень сложно детектировать при тестировании СИИ и проведении клинических испытаний, но которые могут проявить себя при применении СИИ врачами и пациентами.

Методические подходы, общепринятые в доказательной медицине, могут не соблюдаться разработчиками при выводе СИИ на рынок. Например – проведение клинических испытаний ИИ-систем в ретроспективном дизайне исследований может приводить к неполноценной оценке риска. Более того, чаще всего разработчики тестируют свои ИИ-алгоритмы на тех же наборах, что были использованы для машинного обучения, получая за счет этого довольно высокие показатели точности. Это приводит к тому, что при эксплуатации СИИ в реальных условиях существует более высокий риск ошибки при интерпретации данных, чем это было бы возможно без использования технологий ИИ.

## **2.2. Высокий риск деградации метрик точности СИИ при ее применении в реальной клинической практике**

Проведение Московского эксперимента по применению технологий компьютерного зрения в рамках радиологической службы города, а также ряд других публикаций указывают на то, что ИИ-системы склонны демонстрировать достаточно высокую вероятность деградации показателей чувствительности и специфичности во время их эксплуатации в реальной клинической практике [28]. Отчасти это связано с некачественными или неполными наборами данных, использованными при машинном обучении, а также наличием систематических ошибок в таких наборах. Еще одна причина проблемы – изменения в организации работы медицинских организаций, включая появление новых лекарств и медицинского оборудования, а также изменения в течении

заболеваний, из-за чего ИИ-алгоритмы постепенно в рамках своего жизненного цикла начинают терять прежнюю точность работы.

### **2.3. Проблема ответственности за ошибки в работе СИИ**

Любая ИИ-система не может давать гарантированно точный ответ. Всегда есть вероятность ошибки при интерпретации медицинских данных, например, из-за того, что какие-то особые случаи или редкие заболевания не были представлены в обучающем наборе данных, который был использован для создания ИИ-системы. Таким образом, кроме действительно положительного эффекта применения ИИ в практике, есть вероятность дать неправильную подсказку врачу, что в свою очередь может привести к возникновению врачебных ошибок при диагностике и лечении, а значит – причинении вреда здоровью пациента. В настоящее время ответственность за такой вред будет нести тот медицинский работник, который использовал ИИ-систему. По мнению некоторых представителей профессионального сообщества, это является несправедливым. Идет обсуждение предложения о вовлечении разработчиков ИИ-систем в несение ответственности за ошибки в работе их продуктов.

### **2.4. Проблема «черного ящика»**

Ряд методов машинного обучения, в частности – искусственные нейронные сети и глубокое машинное обучение, часто позволяют создавать высокоточные ИИ-алгоритмы. Однако в силу методологических особенностей такие алгоритмы не могут предоставить объяснение полученных выводов. Такое свойство СИИ называется неинтерпретируемостью или, чаще, проблемой «черного ящика». В результате внедрения таких алгоритмов в СИИ для практического применения мы сталкиваемся с ситуацией, когда врач не понимает, почему ИИ-система выдает ему то или иное заключение. Врач начинает относиться к этому с недоверием и в итоге не учитывает или даже отказывается от применения такой системы.

### **2.5. Страхи перед СИИ**

Врачи и руководители в сфере здравоохранения могут относиться к СИИ с подозрением относительно истинных целей их создания и внедрения, а также скрытых особенностей работы. Например, среди врачей может существовать предвзятое мнение, что СИИ могут не только анализировать электронные медицинские карты во время приема и сообщать врачам о возможных врачебных ошибках, но и предоставлять статистику таких ошибок вышестоящим органам, дискредитируя тем самым работу врача в глазах надзорной

инстанции. Еще один из распространенных страхов среди медицинских работников связан с тем, что в будущем ИИ-системы могут заменить их, оставив без работы.

## **2.6. Конфиденциальность медицинской информации**

Для работы СИИ достаточно собирать и анализировать обезличенные медицинские данные. Вместе с этим в среде специалистов по информационной безопасности бытует мнение, что абсолютно надежных методов обезличивания медицинских данных не существует и всегда есть пусть небольшой, но все же риск преобразования обезличенных данных в персональные данные. Для работы СИИ, а также для их обучения, принципиально важным является применение как можно большего объема информации, включая так называемые «большие данные». В этой ситуации пусть даже небольшой риск утечки персональных данных на больших данных создает оправданные опасения по поводу безопасности и конфиденциальности.

## **3. РЕКОМЕНДАЦИИ ПО ЭТИЧЕСКОЙ ЭКСПЕРТИЗЕ КЛИНИЧЕСКИХ ИСПЫТАНИЙ СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА**

Этическая экспертиза испытаний программы клинических испытаний СИИ должна основываться на общих и специфических требованиях к этическим аспектам клинических испытаний.

Общие требования касаются клинических испытаний любых медицинских технологий для профилактики, лечения, диагностики и скрининга:

- обоснованность разработки,
- адекватность дизайна решаемой задаче и обеспечение высокого методологического качества клинических испытаний.

К специфическим требованиям для клинических испытаний СИИ можно отнести:

- зрелость компании-разработчика,
- доверие к наборам данных, использованным для машинного обучения,
- доверие к моделям машинного обучения.

Мы рекомендуем проводить экспертизу по следующему алгоритму (рис. 2):



*Рисунок. 2. Алгоритм проведения экспертизы программы клинических испытаний системы искусственного интеллекта для здравоохранения*

### 3.1. Оценка обоснованности разработки

В технических условиях создания СИИ должна быть представлена цель создания системы, а также научное обоснование и описание ее практической ценности.

Для каждой СИИ должна быть обоснована актуальность рассматриваемой задачи ее создания: почему потребовалась разработка данной системы — на фоне фактического уровня развития здравоохранения, иных исследований и разработок аналогичной тематики.

Производитель СИИ должен определить и теоретически обосновать концепцию или технологию, которая была применена для решения заявленной проблемы и обозначить перспективность такого подхода. Должны быть описаны подходы, которые применялись иными исследователями/разработчиками, а также оценка, почему полученные результаты неудовлетворительны.

Обоснование создания СИИ может быть подкреплено систематическим обзором. При подготовке систематического обзора должны использоваться систематический поиск и систематический отбор исследований. Далее может быть (при наличии оснований) выполнен мета-анализ – количественное обобщение результатов разных исследований. В этом случае возникнет высокая этическая обоснованность для формулировки гипотезы клинического испытания СИИ.

Раскрытие степени исследованности проблемы необходимо для понимания текущего состояния предметной области, обоснования актуальности СИИ, планирования клинических испытаний. Высокий уровень доверия к данной информации может быть обеспечен публикацией производителем СИИ систематического обзора в рецензируемой научной литературе.

#### **Пример № 1**

Пациенты с COVID-19 демонстрируют различную симптоматику, что затрудняет их сортировку. В то время как у некоторых инфицированных пациентов симптомы отсутствуют, у других развивается тяжелый острый респираторный синдром или полиорганная недостаточность. Выявление ключевых характеристик, которые определяют течение болезни в больших когортах пациентов, является актуальной задачей, особенно с учетом его потенциала для помощи врачам и больницам в прогнозировании течения заболевания, эффективном распределении основных ресурсов и улучшении результатов лечения пациентов. Перспективным подходом к решению данной задачи выглядит прогнозирование на основе методов машинного обучения.

## **Пример № 2**

Целенаправленные реабилитационные вмешательства, включая индуцированную ограничением двигательную терапию, двустороннюю тренировку рук, роботизированную терапию и зеркальную терапию, широко используются для устранения двигательной дисфункции у пациентов с последствиями инсульта. Систематические обзоры и метаанализы показали, что эти современные вмешательства эффективны в улучшении двигательной функции у таких пациентов и должны широко применяться в практике. Машинное обучение может стать полезным инструментом для прогнозирования восстановления двигательных функций в процессе реабилитации путем выявления наиболее восприимчивых пациентов, облегчив тем самым практическое использование методик восстановительного лечения.

### **3.2. Оценка дизайна и методологического качества клинического испытания**

Проведение этической экспертизы клинических испытаний СИИ целесообразно проводить, применяя общепринятые подходы в соответствии с правилами надлежащей клинической практики [29] и рекомендациями экспертов [30, 31, 32, 33, 34].

С позиций доказательной медицины клиническое испытание может считаться этически обоснованным в случае выполнения следующих требований:

1. Дизайн исследования соответствует решаемой задаче,
2. Предприняты все необходимые усилия для обеспечения высокого методологического качества испытания, что подразумевает:
  - a. минимизацию систематических ошибок,
  - b. обеспечение правильного статистического сопровождения исследования на всех его этапах.

Поскольку разрабатываемые СИИ решают в большинстве случаев одну из двух задач – диагностики или прогнозирования – то оптимальными дизайнами в этих случаях будут следующие:

- при решении задачи диагностики – формирование одной выборки из целевой популяции; одномоментное независимое применение к каждому пациенту двух тестов - испытываемого и референсного; взаимное маскирование результатов применения этих тестов; случайный порядок их применения; оценка безопасности;



- при решении задачи прогнозирования – формирование одной выборки из целевой популяции; случайное распределение пациентов выборки по двум группам, в одной из которых применяется предлагаемая система прогнозирования и соответствующее прогнозу (возможно - персонализированное) ведение пациента, в другой – рутинное ведение пациентов; оценка клинически значимых (а не суррогатных) исходов; оценка безопасности.

Для минимизации систематических ошибок следует руководствоваться следующими основными документами:

- при решении задач диагностики – STARD, TRIPOD,
- при решении задач прогнозирования – CONSORT.

Указанные документы, как и их варианты для различных ситуаций, приведены на ресурсе <https://www.equator-network.org>.

Правильное статистическое сопровождение исследования включает формулировку гипотезы клинического испытания (превосходства, меньшей эффективности или эквивалентности), расчет минимально необходимого размера выборки, оптимизацию признаков пространства, мониторинг сбора данных, подготовку данных к анализу, корректный анализ данных, правильную интерпретацию и визуализацию его результатов, подготовку прозрачной и добросовестной публикации.

### **3.3. Оценка уровня зрелости компании-разработчика**

Обязательным условием для формирования доверия к компании-разработчику является наличие у нее внедренной системы менеджмента качества (СМК). Самым простым и достоверным способом подтверждения наличия СМК является запрос сертификата добровольной сертификации на соответствие стандарту ISO 13485, выданный независимой организацией-аудитором.

Приветствуется, если разработчик СИИ предоставит свидетельства периодического прохождения независимого аудита СМК, что будет подтверждать не просто однажды внедренной и сертифицированной СМК, но и ее реальное применение и актуализацию у компании-разработчика.

Если программа клинических испытаний подготовлена в целях проведения научного исследования, то наличие СМК у компании-разработчика является добровольным, но не обязательным.

Целесообразно проверить наличие у производителя СИИ свидетельство Роспатента о регистрации на программные продукты.

Рекомендуется оценить наличие научных публикаций у сотрудников компании-разработчика СИИ в рецензируемых изданиях по темам, связанным с разработкой или исследованиям СИИ. Наличие таких публикаций, особенно в журналах, рецензируемых в Web of Science или Scopus, является дополнительным положительным фактором, свидетельствующим о наличии у компании необходимых профильных компетенций, а также практическому опыту по организации, участию или проведению качественных научных исследований.

#### **3.4. Оценка доверия к наборам данных, использованным для машинного обучения**

При машинном обучении должны быть использованы только качественные наборы данных, а процесс их создания и подготовки должен быть прозрачен для эксперта, проводящего этическую оценку клинического испытания СИИ. Непрозрачность источников информации, использованных для создания наборов данных, способов и принципов их формирования следует воспринимать как высокий риск наличия в них систематических отклонений, что может привести к появлению большого количества ошибок при эксплуатации СИИ в условиях реальной клинической практики. В связи с этим при получении комплекта технической документации по СИИ на этическую экспертизу рекомендуется самым пристальным образом оценить ряд важных аспектов, связанных с подготовкой наборов данных к машинному обучению.

##### **3.4.1. Оценка источников, использованных для создания наборов данных**

Любые наборы данных формируются из каких-либо цифровых источников информации, в качестве которых могут выступать данные реальной клинической практики (базы данных электронных медицинских карт, архивы медицинских изображений, данные носимых устройств и т.д.). Еще одним источником информации для создания наборов данных могут выступать регистры или базы данных с результатами клинических исследований. В любом случае производитель СИИ должен предоставить описание источников информации, на основе которых были сформированы наборы данных для обучения и тестирования. В частности, должны быть указаны названия, типы и принадлежность информационных систем, регистров в сфере здравоохранения; общие характеристики содержащихся в них данных; условия и ограничения доступа к данным. Должны быть описаны условия, в которых формировались исходные данные, в том числе временной период, географический регион, виды, формы и условия оказания медицинской помощи и т.д.

### **Пример № 3**

Создание набора данных для машинного обучения было осуществлено на базе регистра Японской ассоциации неотложной медицины догоспитальной остановки сердца (JAAM-ОНСА). Этот регистр представляет из себя проспективную общенациональную базу данных, созданную комитетом реестра JAAM-ОНСА, детально описанную Kitamura T. et al. 2018 и Irisawa T. et al. 2018. Реестр включает догоспитальную, внутрибольничную информацию и сведения об исходах у пациентов с догоспитальной остановкой сердца, доставленных в отделение неотложной помощи. Данные из отделений неотложной помощи собраны в 87 учреждениях (66 университетских больниц и/или третичных центров интенсивной терапии) в Японии. Центры интенсивной терапии сертифицированы Министерством здравоохранения, труда и социального обеспечения Японии и проводят узкоспециализированное лечение, в частности чрескожное коронарное вмешательство. Из включенных в исследование 21 медицинское учреждение не сертифицировано как центр интенсивной терапии, но эти учреждения предоставляют неотложную медицинскую помощь населению, а некоторые могут обеспечить экстракорпоральную сердечно-легочную реанимацию и интенсивную терапию. Всего в реестре JAAM-ОНСА зарегистрировано 34 754 пациента с догоспитальной остановкой сердца с июня 2014 г. по декабрь 2017 г., что является достаточным по мнению разработчиков для проведения машинного обучения.

#### **3.4.2. Оценка наборов данных**

Для каждого использованного при разработке СИИ набора данных должно быть представлено описание процедуры его создания, включая критерии включения / исключения, размерности наборов данных, методы анализа (Data collection/Study Data).

Необходимо оценить, насколько выборки пациентов, присутствующие в наборах данных, соответствуют условиям применения СИИ (целевой популяции). Важно убедиться в том, что СИИ обучалась на той же популяции, для которой она предназначена по техническим условиям.

Также следует убедиться, что набор данных, использованный для машинного обучения, содержит строго обезличенные данные. Персональные данные пациентов не должны присутствовать в наборах данных для обучения, тестирования и валидации моделей машинного обучения.

В технической документации на СИИ должно быть представлено описание пациентов, использованных для создания набора данных (Study population/Participants/Patient Populations и т.п.). Анализируя документацию на СИИ, эксперту необходимо оценить следующие моменты:

1. Характер и объем общей выборки, то есть группы субъектов (пациентов), включаемых в набор данных для создания модели машинного обучения (ММО). Должны быть описаны этническая, национальная, половозрастная, нозологическая, социальная и иные релевантные структуры общей выборки. Если СИИ предназначена для применения в условиях определенной популяции, то и наборы данных должны быть эквивалентны структуре этой популяции и распространенности в ней целевой нозологии.

2. Критерии включения, невключения и исключения. Критерии включения определяются клинической или иной задачей, которая ставится перед разрабатываемой СИИ, чаще они связаны с неким вмешательством — конкретным видом диагностического исследования, хирургического лечения или с конкретной нозологией (патологическим состоянием). Критерии исключения чаще определяют в ходе работы с первичными данными; в целом критерии могут носить медицинский (например, диапазон возраста, наличие сопутствующей патологии и т.д.) и/или технический (параметры и протоколы проведенных диагностических исследований) характер.

3. Механизм обеспечения этики и законности использования данных, включая информацию о соглашении с операторами информационных систем и биомедицинских данных. Опционально, в случае проспективного формирования набора данных, могут рассматриваться аспекты получения информированного добровольного согласия пациента. Должен быть исключен риск, что производитель СИИ получил данные для машинного обучения незаконным способом.

#### **Пример № 4**

В первоначальный набор данных были включены 423 604 пациента, для которых доступно достаточное количество информации. Средний возраст пациентов на исходном уровне составлял 56,4 (стандартное отклонение 8,1) года, и 188 577 участников (44,5%) — мужчины. За период наблюдения (медиана 7 лет) зарегистрировано 6 703 случая сердечно-сосудистых заболеваний (ССЗ). Средний возраст пациентов с ССЗ составил 60,5 года (60,2 года для мужчин и 61,1 года для женщин). Поскольку минимальный период наблюдения для всех участников составлял 5 лет, оценивалась точность

различных моделей в предварительном определении 5-летнего риска развития ССЗ. На 5-летнем горизонте общее количество случаев ССЗ составило 4 801.

Должны быть оценены этапы предварительной обработки данных (препроцессинга), включая очистку и преобразование данных. Необходимо изучить, какие были применены методы работы с пропущенными данными, критерии определения избыточных данных, очистки нереалистичных значений, выбросов и шума (критерии для удаления данных), нормализации, дискретизации, статистической обработки и т.д. По всем примененным методам предварительной обработки данных должна быть предоставлена подробная информация.

#### **Пример № 5**

Для предварительной обработки данных категориальные переменные преобразованы в числовые с однократным кодированием, за исключением параметра «участок окклюзии», которому присваивался порядковый номер (общая сонная артерия — 0, внутренняя сонная артерия — 1 и средняя мозговая артерия сегмент М1 — 2, сегмент М2 — 3 и сегмент М3 — 4). Параметры, имеющие размерность временной отметки, преобразованы в продолжительность времени между наблюдаемыми событиями. Непрерывные переменные дихотомизированы следующим образом: преморбидные по шкале mRS разделены на 0—2 (ранее не нуждавшиеся в уходе) и от 3 до 5 (ранее нуждавшиеся в уходе), поскольку информация о том, является ли пациент нуждающимся в уходе до начала инсульта, имеет решающее значение для прогноза развития событий после инсульта. Время прибытия в больницу разделено на промежутки с 7:00 до 17:00 (прибыли в рабочее время) и с 17:00 до 7:00 (прибыли в нерабочее время), потому что качество лечения пациентов с окклюзией крупных сосудов в острой фазе может заметно различаться в эти периоды. Заполнение пропущенных значений не проводилось. Кроме того, не стандартизировали каждую переменную на этапе препроцессинга, потому что стандартизация со всем набором данных дает дополнительную информацию к каждой переменной. Каждая переменная стандартизирована только после того, как набор данных разделен на обучающие и тестовые данные.

Кроме ошибок в источнике данных (а значит – и в полученных с его помощью наборах данных), еще одной высоковероятной причиной появления ошибок в работе СИИ

может стать этап разметки (аннотирования) набора данных перед обучением. В этой связи важно изучить описание процедуры разметки (аннотирования) наборов данных, включая использованные инструменты, квалификацию персонала и т.д. При этом рекомендуется придерживаться стандартной классификации качества разметки:

- низкая (факт наличия находки) — формируется на основе медицинской документации;
- средняя (классификация находок) — формируется на основе экспертного мнения;
- высокая (подтвержденный диагноз) — формируется на основе результатов более чувствительного метода исследования либо динамического наблюдения (повторного проведения того же метода через определенный временной промежуток).

Необходимо оценить информацию – какие именно эксперты и какого уровня компетенций участвовали в разметке (аннотировании). Целесообразно, чтобы такая разметка осуществлялась несколькими специалистами независимо. Недопустимо, когда уровень компетенций специалистов, участвующих в разметке, будет ниже, чем уровень компетенций медицинских работников, для которых предназначена СИИ.

Целесообразно оценить, использовал ли производитель СИИ имеющиеся стандарты для создания и разметки (аннотирования) наборов данных. Если он их использовал, необходимо проверить – присутствует ли список использованных стандартов в технической документации на СИИ.

В итоговом описании на СИИ должны быть представлены окончательные характеристики включенных в наборы данных пациентов (базовые демографические признаки, клинические признаки, доступные предикторы), в том числе количество лиц (случаев) с отсутствующими, неполными данными и результатом. Практика представления этих результатов в табличном виде с развернутым комментарием выглядит оптимальной. В случае большого объема рекомендуется вынести наглядный материал в приложение. В качестве характеристик набора данных обязательно должны быть указаны:

- общий объем;
- классы, баланс классов, количество записей в каждом классе;
- наличие и вид разметки данных;
- переменные-предикторы;
- стандарты предоставления медицинских данных (например, DICOM).

#### **Пример № 6**

Набор данных для валидации ММО был сформирован в 2019 г. в НПКЦ ДиТ ДЗМ с использованием базы данных результатов программы скрининга рака легкого методом низкодозной КТ (НДКТ). Отбор случаев из базы проведен случайным образом. В набор включены данные 100 лиц, прошедших скрининговое обследование и подписавших информированное добровольное согласие в городских поликлиниках Москвы. Всего было 80 мужчин (средний возраст  $64,8 \pm 9,4$  года) и 20 женщин (средний возраст  $65,9 \pm 8,3$  года); лица европеоидной расы, жители Москвы. Набор данных разделен на 2 подгруппы: «норма» ( $n=32$ ) и «патология» ( $n=68$ ). Общие критерии для подгрупп «норма» и «патология»:

- лица от 55 до 80 лет;
- индекс курения — более 30 пачка/лет;
- курение или отказ от курения в пределах 15 лет;
- отсутствие симптомов (кроме кашля, мокроты, одышки, связанных с курением);
- отсутствие КТ органов грудной клетки в течение последнего года;
- отсутствие в течение последнего месяца повышения температуры тела более  $37,5$  °С, кровохарканья, розовой или «ржавой» мокроты, боли в грудной клетке, сиплого голоса, необъяснимого похудения;
- отсутствие рака легкого и метастазов в легких.

Критерии отнесения к подгруппе «норма»:

- выполнена первичная НДКТ из программы скрининга рака легкого;
- после первичного исследования имеется повторная скрининговая НДКТ, выполненная более чем через 600 дней после первичного исследования;
- после первичного и повторного скрининговых исследований у пациента не выявлен рак легкого.

Критерии отнесения к подгруппе «патология»:

- в течение 3 мес после первичной НДКТ для скрининга верифицировано злокачественное новообразование легкого.

Всем обследуемым выполнена НДКТ грудной клетки. Исследования проведены на 10 компьютерных томографах, имеющих 64 среза Toshiba Aquilion 64 и Toshiba CLX. <...> Данные сканирования обрабатывали, анализировали и сохраняли с помощью программного обеспечения в составе Единого радиологического информационного сервиса.

При использовании наборов данных, созданных иными авторами и размещенных в открытом доступе, как для целей создания моделей, так и для целей их валидации, необходимо корректно процитировать такие наборы.

#### **Пример № 7**

На сайте Московского эксперимента по компьютерному зрению в лучевой диагностике (mosmed.ai) размещен в открытом доступе набор данных с результатами КТ пациентов с COVID-19. В справочном файле представлена информация об условиях использования («Набор данных доступен под лицензией Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0) License») и строка для цитирования в научных и иных публикациях в случае использования этого набора данных сторонними разработчиками.

Приветствуется, если производитель в технической документации на СИИ для каждого примененного набора данных в итоге приведет схему создания набора (рис. 3).





*Рисунок 3. Процесс подготовки данных и проведения машинного обучения с целью создания модели (алгоритма) искусственного интеллекта*

### 3.5. Оценка доверия к моделям машинного обучения

Системы искусственного интеллекта, предназначенные для применения в условиях реальной клинической практики, должны обладать точностью предлагаемых решений на уровне высококвалифицированного эксперта. Как правило, все современные СИИ используют модели машинного обучения (ММО) для интерпретации поступающих в них данных. Именно поэтому проверка корректности и точности работы моделей машинного обучения имеет ключевое значение при проведении клинических испытаний СИИ.

Для оценки уровня доверия к СИИ необходимо оценить прозрачность процесса создания моделей машинного обучения и алгоритмов на основе знаний, включая оценку уровня технологической зрелости. Если информация о том, как были созданы СИИ,

полностью отсутствует в технической документации на нее, такую систему не целесообразно допускать до применения в клинических исследованиях (испытаниях).

Описание процессов машинного обучения должны быть доступны для всех моделей машинного обучения, встроенных в СИИ (Model Development/Model Derivation and Validation/Experimental Evaluation).

Производитель СИИ должен привести обоснование выбора использованных алгоритмов машинного обучения, описание работ по построению модели, методику валидации и метрики оценки качества модели. Должно быть описано, как был разделен набор данных на обучающие и тестовые выборки.

Обратим внимание, что значительное повышение доверия к СИИ обеспечивает проведение производителем независимой внешней валидации моделей на наборах данных, которые не были доступны производителю и не были использованы для обучения (не «оценочный» набор, а полностью новый, например, взятый из сторонних открытых источников), как максимум — валидация независимой группой экспертов. Если внешняя валидация не проводилась, то об этом должно быть заявлено в документации.

Оценка модели должны быть выполнена с учетом правил биостатистики и принципов доказательной медицины, в том числе в части определения объема выборки, выбора методов статистического анализа, оценки конкретных метрик (например, прогностических ценностей положительного и отрицательного результатов, согласованности классификаций и т.д.), способа подтверждения значимости полученных результатов ( $p$ -значение, 95% ДИ). Настоятельно рекомендуется использовать метрики и статистические методы, широко применяемые в медицинской научной литературе, описанные в руководствах по медицинской и биологической статистике.

#### **Пример № 8**

Для разработки моделей прогнозирования использованы два алгоритма машинного обучения:  $k$ -nearest neighbors (KNN) и искусственная нейронная сеть (ANN).

Алгоритм KNN — один из наиболее широко используемых инструментов интеллектуального анализа данных для выявления неизвестных закономерностей в медицинских данных [Zhu M et al., 2007]. Алгоритм KNN предсказывает, что похожие объекты будут существовать в непосредственной близости; в результате он маркирует класс цели на основе окружающих ее  $k$  соседей [Cover T et al., 1967]. Этот метод прогнозирования похож на процесс принятия клинического решения

клиницистами/терапевтами. В большинстве случаев клиницист/терапевт может порекомендовать конкретное вмешательство новому пациенту, если профиль этого нового пациента совпадает с профилями тех пациентов, которые успешно вылечены этим конкретным вмешательством. Таким образом, алгоритм KNN можно рассматривать как искусственную экспертную систему, которая прогнозирует ответы участников на основе обширного опыта, полученного в ходе обучения.

Алгоритм ANN основан на принципах биологических нейронных сетей [Manning T et al., 2014]. Преимущество алгоритмов ANN заключается в том, что они могут фиксировать сложные нелинейные отношения между входными и выходными переменными посредством вычислений в скрытых слоях, что делает их одним из идеальных инструментов для прогнозирования результатов у пациентов с инсультом [Shahid N et al., 2019]. В этом исследовании использован алгоритм ANN с прямой связью.

Целесообразно оценить – была ли реализована производителем СИИ объяснимость результатов работы ММО в интерфейсе СИИ конечному пользователю. Концепция «объяснимого ИИ» критично важна для сферы здравоохранения: процесс, причины и результаты принятия решения моделью МО должны быть доступны (наглядно представлены) и понятны медицинскому работнику — пользователю СИИ.

Для каждой ММО должны быть указаны все ограничения (Limitations) и допущения полученной модели.

#### **Пример № 9 (продолжение примера №8)**

У созданной ММО имеется несколько ограничений. Во-первых, некоторые пациенты с очень большими гематомами или пребывающие в критическом состоянии не включены в исследование по решению семьи, что могло ограничить возможность обобщения результатов. Во-вторых, несмотря на хорошую характеристику алгоритмов, размер выборки относительно невелик. Исследования с более крупными выборками могут продемонстрировать большую прогностическую эффективность. В-третьих, в эту модель не включили информацию о раннем росте гематомы и распространении отека, которые важны для исхода. Наконец, отсутствует внешняя проверка для определения обобщаемости и исключения институциональной предвзятости.

Приветствуется, если производитель СИИ не только обеспечил присутствие информации об обучении в своей технической документации, но и опубликовал результаты создания / тестирования / валидации моделей в рецензируемой научной литературе – такой шаг следует воспринимать как существенное повешение доверия к такой СИИ.

## **Заключение**

Проведение этической экспертизы программы клинических испытаний системы искусственного интеллекта является важнейшим этапом системного формирования доверия к таким продуктам со стороны руководителей в сфере здравоохранения, врачей и специалистов контролирующих органов. Несмотря на то, что практическое внедрение различных СИИ в здравоохранении Российской Федерации началось относительно недавно, тем не менее мы уже имеем достаточное количество готовых решений и различных проектов и, вероятнее всего, их число будет только увеличиваться на фоне повышенного интереса к данным технологиями. В этой связи сотрудникам этических комитетов целесообразно поддерживать направляемые на экспертизу программы испытаний СИИ, детально их рассматривать на основе представленных в данных методических рекомендациях критериях и подходах, внося свой посильный вклад в развитие цифровой трансформации здравоохранения.

## СПИСОК ЛИТЕРАТУРЫ

1. Указ Президента РФ от 10.10.2019 № 490 «О развитии искусственного интеллекта в Российской Федерации» (вместе с «Национальной стратегией развития искусственного интеллекта на период до 2030 года»). Режим доступа: <http://publication.pravo.gov.ru/Document/View/0001201910110003>
2. Тополь Э. Искусственный интеллект в медицине: как умные технологии меняют подход к лечению. – Пер.с англ. – М.: Интеллектуальная литература, 2021. – 434 с. ISBN 978-5-9614-2920-6.
3. Companies in the artificial intelligence in healthcare market are introducing AI-powered surgical robots to improve precision as per the business research company's artificial intelligence in healthcare global market report 2022 [Internet]. Режим доступа: <https://www.globenewswire.com/news-release/2022/03/30/2413072/0/en/Companies-In-The-Artificial-Intelligence-In-Healthcare-Market-Are-Introducing-AI-Powered-Surgical-Robots-To-Improve-Precision-As-Per-The-Business-Research-Company-s-Artificial-Inte.html>
4. State of AI 2021 Report [Internet]. Режим доступа: <https://www.cbinsights.com/research/report/ai-trends-2021/>
5. Альманах Искусственный интеллект. ИИ в здравоохранении. 2022;11 [Internet]. Режим доступа: <https://aireport.ru/healthcare>
6. Basu K., Sinha R., Ong A., Basu T. Artificial Intelligence: How is It Changing Medical Sciences and Its Future? // Indian J Dermatol. 2020 Sep-Oct; 65(5): 365–370. doi:10.4103/ijd.IJD\_421\_20.
7. Кобринский Б.А. Системы искусственного интеллекта в медицинской практике: состояние и перспективы. – Вестник Росздравнадзора. – № 3, 2020. – С. 37–43.
8. Финн В.К. Искусственный интеллект: Методология, применения, философия. – М.: ЛЕНАНД, 2021. – 468 с. ISBN 978-5-9710-6266-0.
9. Esteva A., Chou K., Yeung S. et al. Deep learning-enabled medical computer vision npj Digital Medicine (2021) 4: Article number 5. doi:10.1038/s41746-020-00376-2.
10. Пугачев П.С., Гусев А.В., Кобякова О.С., Кадыров Ф.Н., Гаврилов Д.В., Новицкий Р.Э., Владзимирский А.В. Мировые тренды цифровой трансформации отрасли здравоохранения. Национальное здравоохранение. 2021;2(2):5-12. <https://doi.org/10.47093/2713-069X.2021.2.2.5-12>
11. Topol E. J. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;(25):44–56. doi:10.1038/s41591-018-0300-7.

12. Verghese A., Shah N.H., Harrington R.A. What this computer needs is a physician: humanism and artificial intelligence. JAMA. 2018;(319):19–20. doi:10.1001/jama.2017.19198.
13. Briganti, G., Le Moine O. Artificial Intelligence in Medicine: Today and Tomorrow // Front. Med. 2020;(7):Article 27. doi.org/10.3389/fmed.2020.00027.
14. Price W.N., Gerke S., Cohen I.G. Potential liability for physicians using artificial intelligence. JAMA. 2019;(322):1765–6. doi:10.1001/jama.2019.15064.
15. Кобринский Б.А. Искусственный интеллект в медицине: состояние и горячие точки // Девятнадцатая Национальная конференция по искусственному интеллекту с международным участием КИИ-2021 (11–16 окт. 2021 г.): Тр. конф. / под ред. В.В. Борисова, Б.А. Кобринского. – Ростов-на-Дону; Таганрог: Изд-во Южного федерального ун-та, 2021. – С.13–29.
16. AI and the potential liability issues arising from use in a clinical setting. Available at: <https://www.digitalhealth.net/2019/08/ai-and-the-potential-liability-issues-arising-from-use-in-a-clinical-setting/>.
17. Informatics association seeks tweaks to FDA’s AI regulatory framework. Available at: <https://www.healthcareitnews.com/ai-powered-healthcare/informatics-association-seeks-tweaks-fda%E2%80%99s-ai-regulatory-framework>.
18. Kundu, S. AI in medicine must be explainable. Nature Medicine 27, 1328 (2021). <https://doi.org/10.1038/s41591-021-01461-z>
19. Falco G., Shneiderman B., Badger J. et al. Governing AI safety through independent audits. Nature Machine Intelligence 3, 566–571 (2021). <https://doi.org/10.1038/s42256-021-00370-7>
20. Реброва, О. Ю. Эффективность систем поддержки принятия врачебных решений: способы и результаты оценки / О. Ю. Реброва // Клиническая и экспериментальная тиреоидология. – 2019. – Т. 15, № 4. – С. 148-155. – DOI 10.14341/ket12377. – EDN XFSSOM.
21. Wynants L., Van Calster B., Collins G. S., Riley R. D., Heinze G., Schuit E. et al. Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal BMJ 2020;(369):m1328 doi:10.1136/bmj.m1328, <https://doi.org/10.1136/bmj.m1328>
22. Roberts M., Driggs D., Thorpe M. et al. Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. Nat Mach Intell 3, 199–217 (2021). <https://doi.org/10.1038/s42256-021-00307-0> .

23. Wong A., Otlles E., Donnelly J.P., Krumm A., McCullough J., DeTroyer-Cooley O., Pestrue J., Phillips M., Konye J., Penozza C., Ghaus M., Singh K. External Validation of a Widely Implemented Proprietary Sepsis Prediction Model in Hospitalized Patients. *JAMA Intern Med.* 2021 Aug 1;181(8):1065-1070. doi: 10.1001/jamainternmed.2021.2626. Erratum in: *JAMA Intern Med.* 2021 Aug 1;181(8):1144. PMID: 34152373; PMCID: PMC8218233, <https://doi.org/10.1001/jamainternmed.2021.2626>.

24. Гусев А.В., Морозов С.П., Кутичев В.А., Новицкий Р.Э. Нормативно-правовое регулирование программного обеспечения для здравоохранения, созданного с применением технологий искусственного интеллекта, в Российской Федерации. *Медицинские технологии. Оценка и выбор.* 2021;(1):36-45.

25. Rebrova O.Yu., Gusev A.V. Sample size calculation for clinical trials of medical decision support systems with binary outcome. *Sovremennye tehnologii v medicine* 2022; 14(3): 6, <https://doi.org/10.17691/stm2022.14.3.01>

26. Guy Martin, Saira Ghafur, Isabella Cingolani, Joshua Symons, Dominic King, Sonal Arora et al. The effects and preventability of 2627 patient safety incidents related to health information technology failures: a retrospective analysis of 10 years of incident reporting in England and Wales. *The Lancet Digital Health*, Volume 1, ISSUE 3, e127-e135, July 01, 2019, [https://doi.org/10.1016/S2589-7500\(19\)30057-3](https://doi.org/10.1016/S2589-7500(19)30057-3)

27. Electronic health records vendor to pay largest criminal fine in Vermont history and a total of \$145 million to resolve criminal and civil investigations. Department of Justice, U.S. Attorney's Office, District of Vermont // <https://www.justice.gov/usao-vt/pr/electronic-health-records-vendor-pay-largest-criminal-fine-vermonthistory-and-total-145>

28. Компьютерное зрение в лучевой диагностике: первый этап К63 Московского эксперимента : Монография. 2-е издание, переработанное и дополненное / Ю. А. Васильев [и д. р.]. — [б. м.] : Издательские решения, 2023. — 376 с. ISBN 978-5-0059-3043-9

29. ГОСТ Р 52379-2005. Национальный стандарт Российской Федерации. Надлежащая клиническая практика (утв. Приказом Ростехрегулирования от 27.09.2005 N 232-ст)

30. Вольская Е.А., Хохлов А.Л., Белоусов Д.Ю. Этическая экспертиза исследований реальной клинической практики // *Реальная клиническая практика: данные и доказательства.* 2022. Т. 2. № 2. С. 41-44.

31. Этическая экспертиза биомедицинских исследований: руководство для комитетов по этике / под общей ред. А.Л. Хохлова. – 3-е изд., перераб. и доп. – М.: Изд-во ОКИ, 2021. – 792 с.

32. Вольская Е.А., Семенова Н.В., Громыко Д.И. Практика этической экспертизы: стандартные требования и процедуры независимого этического комитета. Пособие для членов локальных комитетов / Москва, 2015. 41 с.

33. Этическая экспертиза биомедицинских исследований. Практические рекомендации. Под общей редакцией член-корр. РАМН, проф. Ю.Б. Белоусова. Москва, 2005, 156 с.

34. Мартынов А.И., Ющук Н.Д., Вольская Е.А., Ишмухаметов А.А. Этическая экспертиза клинических исследований (основы методологии и практики). Москва, 2001. 72 с.

35. Столбов А.П. О классификации рисков применения медицинского программного обеспечения в Евразийском экономическом союзе. Врач и информационные технологии, 2019; (3): 22-31.

36. Карпов О. Э., Храмов А. Е. Информационные технологии, вычислительные системы и искусственный интеллект в медицине. – М.: ДПК Пресс, 2022. — 480 с., ил.

37. Гусев А.В., Астапенко Е.М., Иванов И.В., Зарубина Т.В., Кобринский Б.А. Принципы формирования доверия к системам искусственного интеллекта для сферы здравоохранения. Вестник Росздравнадзора. 2022; (2):25-33.