

ГБУЗ «НАУЧНО-ПРАКТИЧЕСКИЙ КЛИНИЧЕСКИЙ ЦЕНТР ДИАГНОСТИКИ И
ТЕЛЕМЕДИЦИНСКИХ ТЕХНОЛОГИЙ ДЕПАРТАМЕНТА ЗДРАВООХРАНЕНИЯ
ГОРОДА МОСКВЫ»

ЛУЧШИЕ ПРАКТИКИ ЛУЧЕВОЙ И ИНСТРУМЕНТАЛЬНОЙ ДИАГНОСТИКИ



КЛИНИЧЕСКИЕ ИСПЫТАНИЯ СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА (ЛУЧЕВАЯ ДИАГНОСТИКА)

Москва
2023



ЦЕНТР ДИАГНОСТИКИ
И ТЕЛЕМЕДИЦИНЫ

**ПРАВИТЕЛЬСТВО МОСКВЫ
ДЕПАРТАМЕНТ ЗДРАВООХРАНЕНИЯ ГОРОДА МОСКВЫ**

СОГЛАСОВАНО

Главный внештатный специалист
Департамента здравоохранения
города Москвы по лучевой
и инструментальной диагностике


_____ Ю. А. Васильев
«28» _____ марта 2023 г.

РЕКОМЕНДОВАНО

Экспертным советом по науке
Департамента здравоохранения
города Москвы №4


_____ 2023 г.

**КЛИНИЧЕСКИЕ ИСПЫТАНИЯ
СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА
(ЛУЧЕВАЯ ДИАГНОСТИКА)**

Методические рекомендации №20

Москва
2023

УДК 615.84+616-073.75
ББК 53.6
К 49

Серия «Лучшие практики лучевой и инструментальной диагностики»

Основана в 2017 году

Организация-разработчик:

Государственное бюджетное учреждение здравоохранения города Москвы «Научно-практический клинический центр диагностики и телемедицинских технологий Департамента здравоохранения города Москвы»

Составители:

Васильев Ю. А. – канд. мед. наук, директор ГБУЗ «НПКЦ ДиТ ДЗМ»

Владимирский А. В. – д-р мед. наук, заместитель директора по научной работе ГБУЗ «НПКЦ ДиТ ДЗМ»

Шарова Д. Е. – руководитель отдела инновационных технологий ГБУЗ «НПКЦ ДиТ ДЗМ»

Зинченко В. В. – начальник сектора клинических и технических испытаний ГБУЗ «НПКЦ ДиТ ДЗМ»

Заюнчковский С. Ю. – эксперт сектора клинических и технических испытаний ГБУЗ «НПКЦ ДиТ ДЗМ»

Ахмад Е. С. – научный сотрудник сектора клинических и технических испытаний ГБУЗ «НПКЦ ДиТ ДЗМ»

Гусев А. В. – канд. техн. наук, эксперт сектора клинических и технических испытаний ГБУЗ «НПКЦ ДиТ ДЗМ»

Арзамасов К. М. – канд. мед. наук, руководитель отдела медицинской информатики, радиомики и радиогеномики ГБУЗ «НПКЦ ДиТ ДЗМ»

К 49 Клинические испытания систем искусственного интеллекта (лучевая диагностика) / сост. Ю. А. Васильев, А. В. Владимирский, Д. Е. Шарова [и др.] // Серия «Лучшие практики лучевой и инструментальной диагностики». – Вып. 113. – 2-е изд., перераб. и доп. – М.: ГБУЗ «НПКЦ ДиТ ДЗМ», 2023. – 40 с.

Рецензенты:

Лебедев Георгий Станиславович – д-р техн. наук, директор Института цифровой медицины Первого Московского государственного медицинского университета им. И. М. Сеченова Минздрава России (Сеченовского университета)

Буренчев Дмитрий Владимирович – д-р мед. наук, заведующий отделением, врач службы лучевой диагностики ГБУЗ «Городская клиническая больница им. А. К. Ерамишанцева ДЗМ»

Методические рекомендации предназначены для обеспечения методологической базы при проведении клинических испытаний, клинической апробации и научной оценки алгоритмов и программного обеспечения, разработанных на основе технологий искусственного интеллекта. Клинические испытания рассматриваются как этап подготовки к государственной регистрации программного решения в качестве медицинского изделия. Предложены подходы к оценке точности и эффективности разработок на основе искусственного интеллекта для лучевой диагностики.

Данные методические рекомендации разработаны в ходе выполнения научно-исследовательской работы
«Научно-методические основы цифровой трансформации службы лучевой диагностики»

*Данный документ является собственностью Департамента здравоохранения города Москвы,
не подлежит тиражированию и распространению без соответствующего разрешения*

СОДЕРЖАНИЕ

Нормативные ссылки	4
Определения	5
Обозначения и сокращения	9
Введение	10
Общие положения	12
Случаи, когда проводятся клинические испытания.....	12
Задачи клинических испытаний.....	13
Проведение клинических испытаний.....	13
Предварительный анализ документации.....	14
Подписание договора о проведении клинических испытаний.....	15
Подготовка проекта программы клинических испытаний.....	15
Подготовка дизайна клинических испытаний.....	16
Подготовка окончательной редакции программы клинических испытаний.....	16
Независимый этический комитет (комитет по этике).....	17
Основные принципы выбора медицинской организации для проведения клинических испытаний со стороны производителя системы искусственного интеллекта.....	17
Подготовка производителя системы искусственного интеллекта к клиническим испытаниям.....	18
Порядок проведения оценки клинических функциональных характеристик медицинского изделия.....	20
Специфические требования к формированию выборки для аналитической и клинической валидации системы искусственного интеллекта.....	22
Порядок проведения оценки клинических функциональных характеристик системы искусственного интеллекта.....	23
Подтверждение достоверной клинической связи.....	24
Клиническая валидация.....	25
Аналитическая валидация.....	25
Особенности клинических испытаний системы искусственного интеллекта.....	26
Набор характеристик для оценки эффективности системы искусственного интеллекта.....	28
Заключение	34
Список использованных источников	35
Приложение А. Рекомендованная форма отчета о предварительных клинико-технических испытаниях системы искусственного интеллекта	36

НОРМАТИВНЫЕ ССЫЛКИ

В настоящем документе использованы ссылки на следующие нормативно-правовые акты и нормативные документы (стандарты):

1. Решение Совета Евразийской экономической комиссии от 12.02.2016 № 29 «О правилах проведения клинических и клинико-лабораторных испытаний (исследований) медицинских изделий».
2. Федеральный закон от 27.07.2006 № 152-ФЗ «О персональных данных».
3. Федеральный закон от 21.11.2011 № 323-ФЗ «Об основах охраны здоровья граждан в Российской Федерации».
4. ГОСТ Р ИСО 14155-2014 «Клинические исследования. Надлежащая клиническая практика».
5. ГОСТ Р ИСО 9000 «Системы менеджмента качества. Основные положения и словарь».
6. ГОСТ Р 56429-2021 «Изделия медицинские. Клиническая оценка».
7. ГОСТ Р 59921.1-2022 «Системы искусственного интеллекта в клинической медицине. Часть 1. Клиническая оценка».
8. ГОСТ Р 59921.5-2022 «Системы искусственного интеллекта в клинической медицине. Часть 5. Требования к структуре и порядку применения набора данных для обучения и тестирования алгоритмов».
9. Указ Президента Российской Федерации от 10.10.2019 № 490 «О развитии искусственного интеллекта в Российской Федерации».
10. Приказ Министерства здравоохранения Российской Федерации от 06.06.2012 № 4н «Об утверждении номенклатурной классификации медицинских изделий (с изменениями на 7 июля 2020 года)».

ОПРЕДЕЛЕНИЯ

В настоящем документе применены нижеследующие термины с соответствующими определениями (ГОСТ Р ИСО 14155-2014):

Аналитическая валидация – подтверждение способности системы искусственного интеллекта точно, воспроизводимо и надежно генерировать предполагаемые результаты вычислений из входных данных.

Аудит – систематическая независимая проверка действий и документов, связанных с клиническими исследованиями, для определения, проводятся ли эти действия и ведутся ли записи данных, анализируется ли точность передаваемых данных в соответствии с Программой клинических испытаний (исследований), стандартными рабочими процедурами и применимыми нормативными требованиями.

Гипотеза – утверждение, подвергающееся проверке, основанное на целях и касающееся безопасности или функциональных характеристик исследуемого медицинского изделия, используемое для разработки клинического испытания, которое может быть подтверждено или опровергнуто на основании результатов клинического испытания и статистических вычислений.

Примечание: основная гипотеза является определяющим фактором безопасности исследуемого медицинского изделия или параметров его функциональных характеристик и обычно используется для вычисления объема выборки. Также могут быть оценены дополнительные гипотезы, касающиеся других задач исследования.

Главный исследователь – квалифицированное лицо, ответственное за проведение клинического испытания в исследовательском центре, руководитель исследовательского коллектива.

Дизайн клинического испытания – методология проведения клинического испытания (исследования) с участием человека в качестве субъекта исследования, с помощью которого планируется получить статистически достоверные доказательства клинической безопасности и эффективности испытуемого (исследуемого) медицинского изделия, включающая в себя использование контрольных групп субъектов испытания (исследования), и (или) распределение субъектов испытаний (исследований) по группам случайным образом, и (или) использование одинарного или двойного слепого метода, а также выбор основных и дополнительных конечных точек.

Законный представитель – физическое или юридическое лицо, которое в соответствии с законодательством государства – члена Евразийского экономического союза (далее – ЕАЭС; Союз) имеет право давать информированное согласие на участие в клиническом испытании (исследовании) от имени потенциального субъекта испытания (исследования).

Информированное согласие – письменный документ, в котором субъект испытания (исследования) или его законный представитель подтверждает добровольное согласие на участие в клинических испытаниях (исследованиях) на основании представленной ему полной информации о клиническом испытании (исследовании).

Искусственный интеллект – комплекс технологических решений, позволяющий имитировать когнитивные функции человека (включая самообучение и поиск решений без заранее заданного алгоритма) и получать при выполнении конкретных задач результаты, сопоставимые, как минимум, с результатами интеллектуальной деятельности человека.

Примечание: комплекс технологических решений включает в себя информационно-коммуникационную инфраструктуру, программное обеспечение – ПО (в том числе то, в котором используются методы машинного обучения), процессы и сервисы по обработке данных и поиску решений.

Исследователь – физическое лицо, ответственное за проведение клинического испытания (исследования) в медицинской организации или другом фактическом месте проведения испытания (исследования). В случае, если испытание проводится коллективом исследователей, руководитель данного коллектива является главным исследователем.

Исследовательский центр – фактическое учреждение или место, в котором проводится клиническое испытание, входящее в утвержденный перечень медицинских организаций, проводящих клинические испытания медицинских изделий на территории государств – членов Союза.

Клинические испытания медицинского изделия (КИ) – любое испытание (исследование) с участием человека в качестве субъекта испытания (исследования), проводимое с целью изучения безопасности и (или) эффективности испытуемого (исследуемого) медицинского изделия и (или) метода диагностики или лечения, связанного с его применением.

Проведение клинического испытания (исследования) системы искусственного интеллекта (СИИ) подразумевает использование в качестве субъекта исследования набора данных, который получен с участием человека путем ретроспективного анализа (ГОСТ Р 59921.1-2022).

Примечание: «клиническое испытание» является синонимом термина «клиническое исследование».

Клинические функциональные характеристики медицинского изделия – это результат применения медицинского изделия или оценка субъекта(ов) клинического исследования в отношении медицинского изделия, связанная с его предусмотренным назначением, при условии правильного применения к соответствующему(им) субъекту(ам) клинического испытания.

Клиническая валидация – подтверждение способности системы искусственного интеллекта выдавать клинически значимые выходные данные, связанные с целевым использованием системы искусственного интеллекта в рамках установленного производителем функционального назначения.

Клиническая связь – научное обоснование соответствия результатов действия системы искусственного интеллекта установленному изготовителем (производителем) ее функциональному назначению.

Примечание: научное обоснование достигается путем поиска в литературных источниках, согласно рекомендациям профессиональных сообществ, на основании новых клинических исследований.

Клинические данные – данные по безопасности и (или) эффективности, полученные при клиническом применении медицинского изделия. Клиническими данными также являются данные по безопасности и (или) эффективности, полученные при клиническом применении медицинских изделий, эквивалентность которых рассматриваемому медицинскому изделию может быть доказана.

Клиническое доказательство эффективности и безопасности медицинского изделия – отчет, подтверждающий клиническую эффективность и безопасность медицинского изделия при его использовании по назначению, установленному производителем, на основе клинических данных.

Комитет по этике – независимый экспертный орган, который рассматривает вопросы обеспечения прав, безопасности и охраны здоровья субъектов испытаний (исследований).

Координатор-исследователь – исследователь, назначенный производителем (его уполномоченным представителем) и отвечающий за координацию работы в ходе многоцентрового клинического испытания (исследования).

Многоцентровое клиническое (клинико-лабораторное) испытание (исследование) – клиническое (клинико-лабораторное) испытание (исследование), которое проводится в двух и более медицинских организациях по единой программе.

Мониторинг клинического (клинико-лабораторного) испытания (исследования) – деятельность, связанная с наблюдением за ходом клинического (клинико-лабораторного) испытания (исследования) для проверки того, что данные исследования проводятся и записываются, а отчет о мониторинге формируется в соответствии с программой клинических (клинико-лабораторных) испытаний (исследований), документированными процедурами.

Назначение медицинского изделия – решение производителя в отношении целевого использования медицинского изделия, основанное на его специфических свойствах, обеспечивающих достижение целей медицинского применения и отраженных в технических характеристиках и инструкциях по применению.

Недостаток изделия – несоответствие медицинского изделия требованиям к его идентификации, качеству, надежности, долговечности, безопасности или функциональным характеристикам.

Примечание: недостатки изделия включают отказы, ошибки эксплуатации и неадекватную маркировку.

Отчет о клиническом (клинико-лабораторном) испытании (исследовании) – письменное описание клинического (клинико-лабораторного) испытания (исследования) медицинского изделия с участием человека (или набора данных) в качестве субъекта, объединяющее клиническое (клинико-лабораторное) и статистическое описания, представление данных и их анализ по установленной форме.

Поправка к программе испытаний (исследований) – письменное описание изменений или формальное разъяснение текста программы, которые затрагивают или могут повлиять на достоверность получаемых результатов и исход клинического испытания (исследования).

Программа клинического испытания (исследования) – документ, в котором установлены обоснование, цели, дизайн клинического (клинико-лабораторного) испытания (исследования) и предполагаемый анализ, методология, мониторинг, проведение и ведение записей о клиническом (клинико-лабораторном) испытании (исследовании).

Система искусственного интеллекта (СИИ) – программное обеспечение, в котором используются технологические решения искусственного интеллекта (ГОСТ Р 59921.1-2022).

Спонсор (заявитель) – юридическое или физическое лицо, являющееся инициатором клинического испытания (исследования) и несущее ответственность за его организацию и (или) финансирование.

Субъект клинического исследования – лицо, принимающее участие в клиническом исследовании.

Примечание: субъектом клинического исследования может быть здоровый доброволец или пациент, а также наборы данных.

Технический файл – документированные данные, подтверждающие соответствие медицинского изделия общим требованиям безопасности и эффективности медицинских изделий, требованиям к их маркировке и эксплуатационной документации на них, утверждаемым Евразийской экономической комиссией (ЕЭК).

Уполномоченный орган – орган государственной власти государства – члена Союза, уполномоченный на осуществление и (или) координацию деятельности в сфере обращения медицинских изделий на территории этого государства.

Эффективность медицинского изделия – способность медицинского изделия соответствовать назначению, установленному производителем.

ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ

В настоящем документе применены следующие обозначения и сокращения:
ЕРИС ЕМИАС – Единый радиологический информационный сервис Единой медицинской информационно-аналитической системы

ЕЭК – Евразийская экономическая комиссия

ИИ – искусственный интеллект

ИТ – информационные технологии

КИ – клинические испытания

МИ – медицинское изделие

МО – медицинская организация

НЭК – Независимый этический комитет

ОЛД – отделение лучевой диагностики

ПКТИ – предварительные клинико-технические испытания

ПО – программное обеспечение

СИИ – система искусственного интеллекта

ЭВМ – электронная вычислительная машина

ВВЕДЕНИЕ

Использование технологий искусственного интеллекта (ИИ) в социальной сфере способствует созданию условий для улучшения уровня жизни населения, в том числе за счет повышения качества услуг в сфере здравоохранения (включая профилактические обследования, диагностику, основанную на анализе изображений, прогнозирование возникновения и развития заболеваний, подбор оптимальных дозировок лекарственных препаратов, сокращение угроз пандемий, автоматизацию и точность хирургических вмешательств).

Искусственный интеллект – это научное направление, в рамках которого изучаются проблемы аппаратного или программного моделирования тех видов человеческой деятельности, которые традиционно считаются интеллектуальными; входит в комплекс компьютерных наук; создаваемые технологии относятся к информационным технологиям. Научной новизны в концепции искусственного интеллекта мало. Автоматизированный анализ медицинской информации – это давно существующий раздел науки и практики. Однако в последние годы наблюдается взрывной рост нового поколения технологий ИИ, он связан со значительным инженерным прогрессом вычислительных возможностей компьютерной техники и математики. Национальная стратегия развития ИИ заложена в Указе Президента Российской Федерации от 10.10.2019 № 490 «О развитии искусственного интеллекта в Российской Федерации».

Искусственный интеллект применим для решения ряда задач в медицинской диагностике (поддержка принятия решений, морфометрия, автоматизация участков производственных процессов, контроль качества и т.д.). Для практического применения системы, созданной на основе технологий ИИ, должны пройти технические и клинические испытания для последующей экспертизы с целью государственной регистрации в качестве медицинского изделия. В соответствии с Федеральным законом от 21.11.2011 № 323-ФЗ «Об основах охраны здоровья граждан в Российской Федерации» медицинскими изделиями являются любые инструменты, аппараты, приборы, оборудование, материалы и прочие изделия, применяемые в медицинских целях отдельно или в сочетании между собой, а также вместе с другими принадлежностями, необходимыми для применения указанных изделий по назначению, включая специальное программное обеспечение, и предназначенные производителем для профилактики, диагностики, лечения и медицинской реабилитации заболеваний, мониторинга состояния организма человека, проведения медицинских исследований, восстановления, замещения, изменения анатомической структуры или физиологических функций организма, предотвращения или прерывания беременности, функциональное

назначение которых не реализуется путем фармакологического, иммунологического, генетического или метаболического воздействия на организм человека.

Программное обеспечение, оказывающее влияние на принятие врачом клинически значимых решений либо формирующее новую клинически значимую информацию, несет потенциальный риск применения и тем самым возможный риск причинения вреда здоровью пациента.

ПО является медицинским изделием (МИ) при условии соответствия совокупности следующих критериев (согласно приказу Министерства здравоохранения Российской Федерации от 06.06.2012 № 4н):

- представляет собой программу для электронной вычислительной машины (ЭВМ) или ее модули вне зависимости от используемой аппаратной платформы, а также способов размещения ПО и предоставления доступа к нему;
- не является составной частью другого МИ;
- предназначено производителем для оказания медицинской помощи;
- результат действия ПО заключается в интерпретации в автоматическом режиме, в том числе с использованием технологий ИИ, или по заданным медицинским работником параметрам, влияющим на принятие клинических решений, набора данных, полученных от МИ, допущенных к обращению в установленном порядке или введенных медицинскими работниками в целях оказания медицинской помощи.

Таким образом, ПО с применением технологий ИИ подлежит государственной регистрации в качестве МИ. Обязательным подготовительным этапом перед направлением такого ПО на регистрацию в качестве МИ является проведение технических и клинических испытаний. Данные методические рекомендации устанавливают единообразные подходы к определению эффективности ПО в процессе клинических испытаний.

Итак, важные свойства ПО с применением технологий ИИ [1]:

- а) ПО используется в качестве самостоятельного МИ;
- б) ПО не является принадлежностью (составной частью) другого МИ;
- в) ПО не предназначено для установки (инсталляции) на этом МИ.

Далее в тексте такое ПО, в котором используются технологические решения ИИ, будет называться «система искусственного интеллекта» (СИИ).

В данных методических рекомендациях будут рассмотрены как общие применяемые требования и рекомендации по проведению клинических испытаний МИ в странах Евразийского экономического союза (ЕАЭС), так и особенности проведения клинических испытаний программного обеспечения (ПО), в том числе с применением технологий ИИ (СИИ). Выполнение клинических испытаний СИИ в рамках национальной системы осуществляется в соответствии с методическими рекомендациями ФГБУ «ВНИИИМТ» Росздравнадзора [2].

ОБЩИЕ ПОЛОЖЕНИЯ

Клинические испытания являются одной из форм оценки соответствия МИ установленным требованиям.

Клинические испытания проводятся в медицинских организациях (МО). Национальный перечень медицинских организаций, имеющих право проводить клинические испытания МИ, доступен на сайте Федеральной службы по надзору в сфере здравоохранения (Росздравнадзор) по адресу: http://www.roszdravnadzor.ru/services/test_clinical. Полный перечень МО, уполномоченных проводить клинические испытания МИ в целях их регистрации в рамках Евразийского экономического союза, представлен в Едином реестре уполномоченных организаций, имеющих право проводить исследования (испытания) медицинских изделий в целях их регистрации, размещенном в открытой части информационной системы Союза в сфере обращения медицинских изделий (https://portal.eaeunion.org/sites/odata/_layouts/15/Registry/PMM07/TableView.aspx).

Соответственно, предоставление результатов клинических испытаний от МО, не включенной в перечень, автоматически приведет к отказу в регистрации МИ.

Клиническое испытание – это один из заключительных этапов в процессе подготовки МИ к государственной регистрации.

Случаи, когда проводятся клинические испытания

Клинические испытания МИ с участием человека проводятся в следующих случаях:

- а) новый вид медицинского изделия;
- б) применение новых сложных, и (или) уникальных, и (или) специальных методов профилактики, диагностики и лечения заболеваний и состояний, а также применение новых сложных медицинских технологий;
- в) при проведении анализа и оценки клинических данных не подтверждены эффективность и безопасность МИ.

В остальных случаях клинические испытания МИ проводятся в форме анализа и оценки клинических данных.

Все клинические исследования проводятся с соблюдением специальных международных правил надлежащей клинической практики – Good Clinical Practice; GCP (ГОСТ Р ИСО 14155-2014).

Соблюдение стандартов GCP служит для общества гарантией того, что права, безопасность и благополучие субъектов исследования защищены, согласуются с принципами, заложенными Хельсинкской декларацией Всемирной медицинской ассоциации [3], и что данные клинического исследования достоверны.

Задачи клинических испытаний

Задачи, решаемые в ходе проведения клинических испытаний (ГОСТ Р 59921.1-2022):

- а) подтверждение соответствия МИ нормативной документации, технической и эксплуатационной документации производителя;
- б) подтверждение соответствия представленной заявителем документации установленным производителем назначению и показаниям к применению;
- в) оценка полноты и достоверности установленных нормативной документацией, технической и эксплуатационной документацией производителя характеристик МИ;
- г) оценка качества МИ, эффективности и безопасности его применения, в том числе ожидаемого терапевтического эффекта в отношении МИ, применяемых для терапии.

При подтверждении соответствия указанным требованиям результаты клинических испытаний считаются положительными.

Результаты клинических испытаний МИ считаются отрицательными в следующих случаях:

- а) медицинское изделие не соответствует назначению и показаниям к применению, установленным производителем в эксплуатационной документации на медицинское изделие;
- б) выявлены побочные действия, не указанные в инструкции по применению или руководстве по эксплуатации медицинского изделия, нежелательные реакции при его применении;
- в) установлены факты и обстоятельства, создающие угрозу жизни и здоровью граждан и медицинских работников при применении и эксплуатации медицинского изделия.

Проведение клинического испытания (исследования) МИ класса потенциального риска применения 3 или 2б (в соответствии с Приказом Министерства здравоохранения Российской Федерации от 06.06.2012 № 4н), а также имплантируемого МИ осуществляется в обязательном порядке в форме многоцентровых испытаний (исследований), в том числе в одном из государств – членов ЕАЭС.

Проведение клинических испытаний

Готовность к проведению клинических испытаний обеспечивается как подготовкой МО, так и подготовкой производителя медицинского изделия (СИИ) к проведению клинических испытаний своего медицинского изделия.

Порядок проведения клинического испытания представлен на рисунке 1.

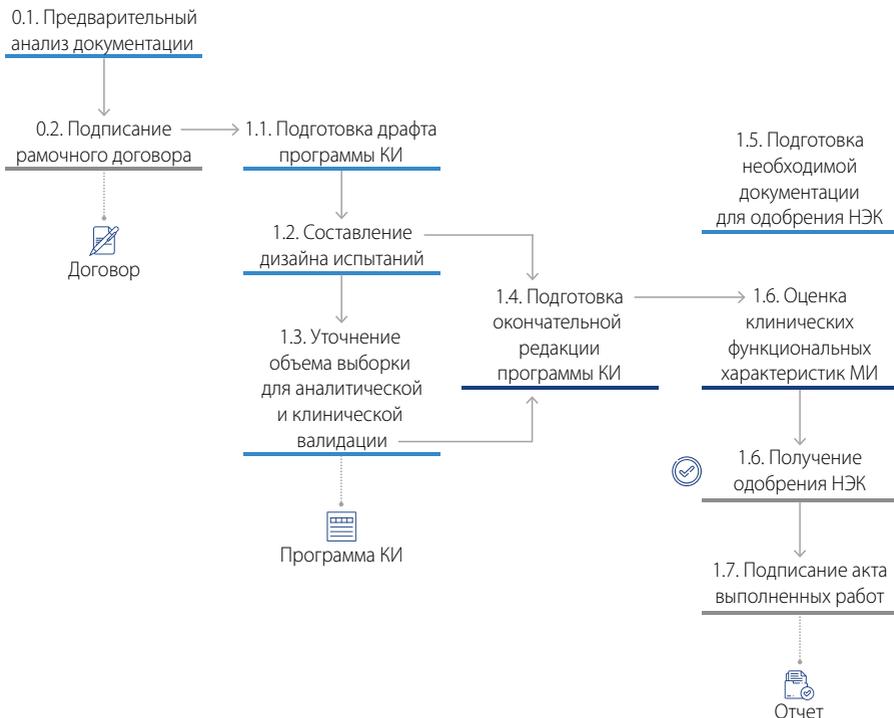


Рисунок 1 – Порядок проведения клинических испытаний

Предварительный анализ документации

Проведение клинических испытаний осуществляется на основе договора между МО и производителем МИ. Для заключения договора о проведении клинического испытания необходим предварительный анализ документации.

Производитель или его уполномоченный представитель направляет в МО следующие документы (в случае, если оригиналы документов составлены на иностранном языке, они представляются с переводом на русский язык, заверенным производителем или его уполномоченным представителем):

- а) заявление на проведение клинического испытания (исследования) МИ;
- б) заявление с подтверждением того, что данное медицинское изделие удовлетворяет всем применимым требованиям безопасности и эффективности, исключая те свойства и характеристики, которые должны быть исследованы в ходе клинических испытаний (исследований) медицинских изделий;
- в) проект программы клинического испытания (исследования);

г) образцы (образец) медицинского изделия вместе с принадлежностями, необходимыми для применения медицинского изделия по назначению (при наличии принадлежностей);

д) сведения о соответствии медицинского изделия общим требованиям безопасности и эффективности медицинских изделий, требованиям к их маркировке и эксплуатационной документации на них;

е) инструкцию по применению (эксплуатационную документацию) медицинского изделия с принадлежностями, необходимыми для применения медицинского изделия по назначению (при наличии принадлежностей);

ж) документы (материалы), содержащие результаты собственных испытаний (исследований) производителя, в том числе анализ данных применения медицинского изделия (при наличии);

з) результаты испытаний в целях утверждения типа средств измерений (для медицинских изделий, включенных в перечень медицинских изделий, относящихся к средствам измерений, в отношении которых проводятся испытания в целях утверждения типа средств измерений);

и) одобрение комитета по этике проведения клинического испытания (исследования) медицинского изделия.

Подписание договора о проведении клинических испытаний

МО готовит предложение по составу исследовательского коллектива для проведения КИ, включающего: главного исследователя, исследователей (соисследователей), исследователя-координатора (в случае многоцентровых исследований). С этой целью проводится оценка уровня квалификации главного исследователя и других участников исследовательского коллектива; изучаются профессиональные автобиографии участников; обеспечивается документальное подтверждение подготовки, опыта и научных или клинических знаний всех вовлеченных сторон; выявляется возможный конфликт интересов между главным исследователем и соисследователями (ГОСТ Р 59921.1-2022, ГОСТ Р 56429). Проводится предварительное согласование состава исследовательского коллектива с его участниками. Исследовательский коллектив утверждается в тексте приказа о проведении КИ. Далее проводится подписание договора, и начинаются работы в рамках клинических испытаний.

Подготовка проекта программы клинических испытаний

Требования к содержанию программы клинических испытаний (исследований) закреплены в Решении Совета Евразийской Экономической комиссии от 12.02.2016 № 29, а также в ГОСТ Р ИСО 14155-2014 «Клинические исследования».

Надлежащая клиническая практика» и ГОСТ Р 56429-2015 «Изделия медицинские. Клиническая оценка».

Программа КИ состоит из двух частей: общего описания медицинского изделия и данных о предстоящих клинических испытаниях (исследовании). Программа КИ разрабатывается в соответствии с описанием назначения медицинского изделия, анализом данных, касающихся эксплуатационных характеристик медицинского изделия, и предполагаемым методом применения. Программа КИ должна содержать общие характеристики МИ и конкретную информацию об организации, ходе клинического испытания, описания его дизайна, целей, гипотез, статистических методов анализа, мониторинге, порядке взаимодействия с комитетом по этике и т.д.

Подготовка дизайна клинических испытаний

Дизайн КИ, являющийся составной частью программы КИ, должен быть разработан таким образом, чтобы полученные результаты испытаний (исследований) позволили оценить, подходит ли исследуемое медицинское изделие для цели(ей) и групп пациентов, для которой(ых) оно предназначено, а также его клиническую эффективность.

Описание дизайна клинического испытания должно включать в себя:

- а) стратегию обеспечения клинической значимости и научной обоснованности результатов с обоснованием ее выбора;
- б) описание цели и гипотезы испытаний;
- в) описание выбора субъектов клинических испытаний (исследований), в том числе их количество;
- г) требования к набору данных;
- д) требования к объему выборки;
- е) основные и дополнительные конечные точки с обоснованием их выбора, и измерения;
- ж) модель и тип исследований (например, межгрупповое, продольное, ретроспективное) и т.д.

Рекомендуется при описании дизайна приводить блок-схему испытания.

Подготовка окончательной редакции программы клинических испытаний

Производитель представляет в МО проект программы КИ.

МО рассматривает проект программы и, в случае необходимости, редактирует ее. Окончательная редакция программы направляется в адрес производителя. Производитель СИИ письменно согласовывает, а МО утверждает программу КИ.

МО приступает к КИ при наличии:

- заключенного договора на проведение КИ;
- утвержденной программы КИ;
- положительного заключения комитета по этике;
- образцов МИ,
- необходимой документации.

Независимый этический комитет (комитет по этике)

Заявитель представляет в комитет по этике заявление и все необходимые документы, регламентированные Решением Совета Евразийской экономической комиссии от 12.02.2016 № 29 «О правилах проведения клинических и клинико-лабораторных испытаний (исследований) медицинских изделий».

Комитет по этике рассматривает вопрос о проведении предлагаемого клинического испытания (исследования) в соответствующие сроки и дает заключение в письменном виде, в котором должно быть идентифицировано испытание (исследование) с указанием рассмотренных документов и даты принятия решения.

КИ не должно начинаться до получения письменного одобрения / положительного решения комитета по этике.

Основные принципы выбора медицинской организации для проведения клинических испытаний со стороны производителя системы искусственного интеллекта

Для проведения клинических испытаний производитель СИИ выбирает МО, с которой он будет заключать договор о проведении КИ.

Перечень МО, имеющих право проводить клинические испытания (исследования) медицинских изделий в целях их регистрации, ограничен. Требования к МО, уполномоченной проводить клинические испытания, содержатся в Решении Совета Евразийской экономической комиссии от 12.02.2016 № 29.

Для того чтобы МО была уполномочена Росздравнадзором на проведение клинических испытаний медицинских изделий в целях их регистрации в рамках Евразийского экономического союза, необходимо подать заявку в Росздравнадзор с необходимым комплектом документов.

Заявка должна содержать информацию о подтверждении требований, указанных в ч. V Правил проведения клинических и клинико-лабораторных испытаний (исследований) медицинских изделий, утвержденных Решением Совета Евразийской экономической комиссии от 12.02.2016 № 29.

Руководитель МО подписывает заявку и сопроводительное письмо, а также заверяет сопутствующий комплект документов:

- копию лицензии на осуществление медицинской деятельности (с указанием перечня работ (услуг), составляющих медицинскую деятельность);
- копию Устава МО;
- копии утвержденного штатного расписания МО;
- копию санитарно-эпидемиологического заключения;
- копию Положения о проведении клинических испытаний;
- копии положений о структурных подразделениях и иных положений МО;
- копии разработанных МО стандартных операционных процедур (такие как «Взаимодействие с комитетом по этике», «Обучение и квалификация исследовательского коллектива», «Доступ к персональным данным и их обработка», «Обеспечение защиты конфиденциальной информации», «Аудит клинического испытания», «Подготовка, обращение и обновление стандартной операционной процедуры», «Получение информированного согласия»);
- копии документов об образовании членов исследовательского коллектива.

Только после положительного решения Росздравнадзора и внесения в реестр МО вправе проводить клинические испытания.

Резюмируя вышеизложенное, можно выделить основные принципы выбора МО для проведения КИ со стороны производителя СИИ:

1. Наличие у МО действующей лицензии на осуществление медицинской деятельности.
2. Наличие МО в национальном перечне Федеральной службы по надзору в сфере здравоохранения (Росздравнадзор) (http://www.roszdravnadzor.ru/services/test_clinical).
3. Наличие МО в Едином реестре уполномоченных организаций, имеющих право проводить исследования (испытания) медицинских изделий в целях их регистрации, размещенном в открытой части информационной системы Союза в сфере обращения медицинских изделий (https://portal.eaeunion.org/sites/odata/_layouts/15/Registry/PMM07/TableView.aspx).
4. Наличие у МО компетентного исследовательского коллектива и исследовательской базы.

Подготовка производителя системы искусственного интеллекта к клиническим испытаниям

Для проведения клинических испытаний производитель СИИ предоставляет в МО следующие документы (Решение Совета ЕЭК № 29):

- а) разрешение уполномоченного органа на проведение клинического испытания (исследования) медицинского изделия;
- б) заявление на проведение клинического испытания (исследования) медицинского изделия;
- в) заявление с подтверждением того, что данное медицинское изделие удовлетворяет всем применимым требованиям безопасности и эффективности, исключая те свойства и характеристики, которые должны быть исследованы в ходе клинических испытаний (исследований) медицинских изделий;
- г) брошюру исследователя;
- д) проект программы клинического испытания (исследования);
- е) образец (образцы) медицинского изделия вместе с принадлежностями, необходимыми для применения медицинского изделия по назначению (при наличии принадлежностей);
- ж) сведения о соответствии медицинского изделия общим требованиям безопасности и эффективности медицинских изделий, требованиям к их маркировке и эксплуатационной документации на них;
- з) инструкции по применению (эксплуатационную документацию) медицинского изделия с принадлежностями, необходимыми для применения медицинского изделия по назначению (при наличии принадлежностей);
- и) материалы, содержащие результаты собственных испытаний (исследований) производителя, в том числе анализ данных применения медицинского изделия (при наличии);
- к) результаты испытаний в целях утверждения типа средств измерений (для медицинских изделий, включенных в перечень медицинских изделий, относящихся к средствам измерений, в отношении которых проводятся испытания для утверждения типа средств измерений);
- л) одобрение комитета по этике на проведение клинического испытания (исследования) медицинского изделия;
- м) копию документа об условиях страхования или возмещения (компенсации) возможного вреда при возникновении неблагоприятных событий (инцидентов), осуществляемых в соответствии с законодательством государства – члена ЕАЭС, на территории которого проводится клиническое испытание (исследование) медицинского изделия;
- н) форму информированного согласия;
- о) форму индивидуальной регистрационной карты;
- п) форму отчета о неблагоприятных событиях (инцидентах);
- р) форму отчета о недостатках медицинского изделия;
- с) форму отчетов о клиническом испытании (исследовании) медицинского изделия, включая промежуточный отчет.

Согласно Решению Совета ЕЭК № 29 также необходимо предоставлять технический файл, содержащий в том числе результаты испытания в условиях, имитирующих эксплуатационные. Для СИИ такие испытания проводят в рамках предварительных клинико-технических испытаний (ПКТИ). Для проведения ПКТИ СИИ производитель должен установить перечень проверяемых параметров, разработать программу испытаний и критерии прохождения [4]. ПКТИ проводят в медицинских организациях, которые могут сформировать наборы данных, необходимые для проведения верификации и валидации СИИ. По результатам оформляют отчет в соответствии с приложением к настоящему методическим рекомендациям.

К основным ошибкам, допускаемым производителем при подготовке СИИ к КИ, можно отнести следующие:

1. Технические и клинические испытания проходят со значительным временным перерывом, что приводит к «устареванию» нормативных требований (стандартов), на соответствие которым была проверена СИИ в ходе технических испытаний. Как следствие, на момент проведения КИ результаты технических испытаний теряют свою актуальность.

2. В представленной документации производителя некорректно определяется класс риска СИИ.

3. В документации производителя отсутствует информация об аналогичных (взаимозаменяемых) медицинских изделиях.

4. Не представляется информация о жизненном цикле СИИ и файл менеджмента рисков.

5. В представленной на клинические испытания технической и эксплуатационной документациях не в полной мере отражена информация о назначении СИИ, отсутствует информация о показаниях и противопоказаниях к медицинскому применению медицинского изделия.

6. В проекте программы КИ некорректно формулируется гипотеза КИ.

Порядок проведения оценки клинических функциональных характеристик медицинского изделия

Клинические испытания МИ должны проводиться в соответствии с принципами Хельсинкской декларации Всемирной медицинской ассоциации 1964 года «Этические принципы медицинских исследований с участием человека в качестве субъекта исследования».

Клинические испытания медицинских изделий должны проводиться в условиях, установленных производителем для применения медицинского изделия, и должны быть предусмотрены программой клинического испытания (исследования).

Порядок проведения оценки клинических функциональных характеристик СИИ может быть представлен в виде последовательного исполнения следующих этапов работ.

Этап 1. Определяются условия проведения испытаний, а именно:

- формируется экспертная группа;
- определяется необходимое техническое оснащение;
- обеспечивается доступ к медицинским данным ответственным лицам;
- решаются организационные вопросы, связанные со взаимодействием с разработчиком, согласовываются даты мониторингования КИ, изучаются внутренние стандарты организации (СОПы) и т.д.;
- осуществляются работы по инсталляции ПО;
- проводится мониторинг, требуемый к проведению в начале испытаний.

Этап 2. Основное внимание уделяется разработке дизайна испытаний:

- подтверждается клиническая связь;
- формируется гипотеза КИ (согласно требованиям к клинической валидации);
- определяются метрики качества;
- формируются требования к объему выборки;
- формируются требования к наборам данных;
- определяются методы анализа данных.

Этап 3. Связан с непосредственным проведением оценки функциональных характеристик МИ, представленного на испытания:

- проводится сбор набора данных;
- выполняется тестирование МИ;
- проводится анализ (расчет метрик, экспертный анализ работы ПО);
- проводятся промежуточные мониторинги испытаний.

Этап 4. Подготовка отчета о результатах исследования:

- определяются содержание и объем отчета клинических испытаний МИ;
- решаются организационные вопросы по взаимодействию с разработчиком в части деинсталляции ПО, завершению испытаний.

Отчет о клиническом испытании (с приложением обобщающих таблиц (графиков) результатов испытаний с соответствующей статистической обработкой и комментариями к ним) должен содержать критическую оценку всех данных, полученных в ходе испытаний, в том числе негативных данных. При проведении многоцентровых клинических испытаний составляется общий отчет, который должен быть подписан исследователями и руководителями медицинских организаций и утвержден координатором-исследователем.

Заявитель имеет право контролировать проведение клинического испытания медицинского изделия на всех этапах посредством назначения лиц соот-

ветствующей квалификации и проведения процедур мониторинга или независимой оценки соответствия с целью получения достоверных данных.

Специфические требования к формированию выборки для аналитической и клинической валидации системы искусственного интеллекта

В программе КИ должны быть четко определены гипотеза и цели КИ (первичные и вторичные) СИИ, а также объем выборки (с расчетами), заявленные характеристики и предполагаемая эффективность изделия, подлежащие верификации [5].

Одним из самых сложных практических вопросов разработки программы КИ является определение размера выборки пациентов, необходимого и достаточного для формирования надежных выводов о качестве работы медицинского изделия. В конечном счете это необходимо для формирования доверия к результатам клинических испытаний медицинского изделия и, в частности, программного продукта, содержащего модели машинного обучения, в процессе регистрации его в качестве медицинского изделия. Таким образом, корректность определения размера выборки может стать важнейшим фактором успешности проведения КИ и последующего получения разработчиком государственного регистрационного удостоверения.

В программе КИ должны быть указаны методы анализа и обработки данных, включая статистические методы обоснования объема выборки (ГОСТ Р 59921.1-2022, ГОСТ Р 59921.5-2022).

Важным нюансом является то, что различные модели машинного обучения, входящие в состав СИИ, могут быть предназначены для разных условий применения и критериев срабатывания, например, в зависимости от условий оказания медицинской помощи, пола и возраста пациента, наличия или отсутствия у него определенных заболеваний и т.д. Таким образом, рекомендуется определять размер выборки валидационного набора данных не в комплексе для всей СИИ, а отдельно для каждой модели машинного обучения.

Также необходимо отметить, что объем выборки является не единственной ее принципиально важной характеристикой – важна также ее репрезентативность. Репрезентативность выборки может быть достигнута при вероятностном способе ее формирования (случайный отбор, систематический отбор, кластерный отбор и т.д.), однако на практике такой подход в большинстве случаев невозможен. Обычно используются следующие невероятностные (неслучайные) способы формирования выборок: выборка удобства (произвольная), последовательная (сплошная) выборка, выборка добровольцев, квотный отбор и т.д. В этой ситуации важно обращать внимание на распространенность (долю,

частоту) диагностируемого (или прогнозируемого) состояния в валидационной выборке, т.к. именно от него зависят наиболее важные операционные характеристики (метрики точности) модели.

Отбор записей для валидационного набора данных следует выполнять из той же популяции пациентов, в которой предполагается применять разработанную СИИ – целевой популяции. Так, если модель, предназначенная для встраивания в программный продукт, разработана на данных пациентов из истории болезни стационара, то и валидировать, и применять данный продукт следует для таких же пациентов, а не для пациентов, например, поликлиники. В идеале и обучающая выборка должна быть репрезентативной хотя бы в отношении распространенности диагностируемого (прогнозируемого) состояния.

Практические подходы к расчету необходимого размера выборки для валидации диагностических и прогностических моделей приведены в публикации [6].

Порядок проведения оценки клинических функциональных характеристик системы искусственного интеллекта

Проведение оценки клинических характеристик СИИ в общем виде осуществляется в следующем порядке:

- анализ и оценка клинических данных документов и материалов, представленных заявителем;
- оценка сведений о проводившихся клинически значимых корректирующих действиях, в том числе о приостановлении применения медицинского изделия, об изъятии из обращения медицинского изделия, об отзывах медицинского изделия;
- анализ научной литературы и (или) неопубликованных данных и сообщений, соотнесенных с предназначенным производителем применением испытываемого медицинского изделия и предлагаемым методом его использования;
- редактирование и согласование программы клинических испытаний, представленной заявителем;
- проведение испытаний образцов (образца) медицинского изделия;
- доработка эксплуатационной документации производителя на медицинское изделие по результатам испытаний (при необходимости);
- оформление и выдача (вручение или направление заказным почтовым отправлением с уведомлением о вручении) заявителю отчета о клиническом испытании (исследовании) медицинского изделия.

Существенная особенность проведения оценки клинических функциональных характеристик СИИ состоит в том, что в качестве субъекта исследования используется набор данных, который получен с участием человека путем ретроспективного анализа и т.п.

В ходе клинических испытаний проводится процесс оценки достоверности клинической связи, аналитической и клинической валидации в соответствии с приказом Министерства здравоохранения Российской Федерации от 30.08.2021 № 885н «Об утверждении Порядка проведения оценки соответствия медицинских изделий в форме технических испытаний, токсикологических исследований, клинических испытаний в целях государственной регистрации медицинских изделий».

Подтверждение достоверной клинической связи

В процессе проведения клинических испытаний должно быть дано обоснование соответствия выходных данных СИИ заявленному функциональному назначению, т.е. должна быть подтверждена достоверная клиническая связь. Клиническая связь подтверждает свойство СИИ генерировать выходные данные, которые коррелируют с клиническими факторами исследуемого объекта. Определение клинической связи выполняют путем анализа имеющихся данных или получения новой информации из литературных источников, документации, проведенных предварительных испытаний и т.д.

Подтверждение достоверности клинической связи означает, что:

- программное обеспечение имеет устойчивую клиническую связь, базирующуюся на документально подтвержденных результатах, как это определено в клинических рекомендациях, клинических исследованиях в рецензируемых журналах, согласии на использование программного обеспечения, международных справочных материалах или других подобных хорошо зарекомендовавших себя данных;

- программное обеспечение имеет новую клиническую значимость, базирующуюся на способности использовать новые входные данные, алгоритмы или результаты, новую целевую группу или новое предназначение.

Подтверждение достоверности клинической связи – необходимый компонент клинической оценки для любого программного обеспечения, являющегося медицинским изделием, и может быть продемонстрировано:

- а) ссылками на существующие данные исследований, проведенных для того же предполагаемого использования;

- б) ссылками на существующие данные исследований, проведенных для другого предполагаемого использования, где может быть оправдана экстраполяция таких данных;

- в) генерацией и определением новых клинических данных для конкретного предполагаемого использования программного обеспечения.

Результатом клинических испытаний на этом этапе является подтверждение достоверной клинической связи предназначения программного обеспечения с его целевым использованием.

Клиническая валидация

В процессе проведения клинических испытаний должна быть подтверждена способность СИИ выдавать клинически значимые выходные данные, связанные с целевым использованием системы искусственного интеллекта в рамках установленного производителем функционального назначения, т.е. должна быть проведена клиническая валидация СИИ.

В процессе проведения клинических испытаний на данном этапе должны быть предоставлены подтверждения, объективные свидетельства того, что требования, устанавливаемые для конкретного использования или применения медицинского изделия, выполнены (ГОСТ Р ИСО 9000).

Для решения этих задач требуется провести инсталляцию (установку) СИИ для обеспечения доступа к ней непосредственного пользователя.

Порядок инсталляции СИИ определяется производителем и отражается в Программе проведения КИ.

В процессе клинической валидации должны быть решены следующие задачи:

1. Идентификация СИИ (наименование, номер и дата версии).
2. Проверка совместимости операционных систем и аппаратных средств, их идентификация (если применимо).
3. Подтверждение того, что СИИ функционирует в полном соответствии с технической и эксплуатационной документацией.
4. Подтверждение, что СИИ не приводит к критическим отказам, в том числе в предельных режимах нагрузки (длительность использования изделия, механические нагрузки на изделие, скорости и мощность работы, вибрационные нагрузки, скачки питающего напряжения и т.д. в зависимости от изделия, нормативных условий его эксплуатации и назначения).
5. Подтверждение, что интерфейс СИИ сводит к минимуму вероятность ввода ошибочных или случайных команд в тех случаях, когда это может пагубно повлиять на лечебный или диагностический процесс.

Результатом клинических испытаний на этом этапе является подтверждение клинической валидации программного обеспечения как медицинского изделия, т.е. соответствие его характеристик, полученных в результате клинических испытаний, характеристикам, указанным в технической и эксплуатационной документации.

Аналитическая валидация

Аналитическая валидация включает:

- подтверждение и предоставление объективных доказательств того, что СИИ была правильно сконструирована, а именно, правильно и надежно

обрабатывает входные данные и генерирует выходные данные с надлежащими уровнем точности и повторяемостью;

- подтверждение того, что СИИ соответствует заявленным требованиям производителя, а также что требования к СИИ соответствуют потребностям пользователя и предполагаемому использованию.

Особенности клинических испытаний системы искусственного интеллекта

Проведение КИ СИИ имеет ряд особенностей. В первую очередь это касается процесса клинической валидации.

Следует иметь в виду, что, как было указано ранее, в испытаниях СИИ субъектом испытаний является набор данных, полученный с участием человека путем ретроспективного анализа. С учетом этого клинические испытания СИИ должны проводиться с соблюдением всех применимых этических требований и норм.

Этические аспекты должны быть в первую очередь учтены при следующих процедурах клинических испытаний СИИ:

- при формировании программы клинических испытаний;
- при оценке соответствия квалификации исследователей предлагаемому испытанию;
- при сборе данных и формировании наборов данных для проведения испытаний;
- при анализе возможных неблагоприятных событий, возникающих при использовании СИИ;
- при анализе документов с результатами клинических испытаний;
- при установке и деинсталляции СИИ для цели клинических испытаний;
- при взаимодействии с заказчиком клинических испытаний в течение всей их продолжительности и др.

Об особенностях и требованиях к *формированию выборки* при проведении КИ СИИ говорилось в разделе «Специфические требования к формированию выборки для аналитической и клинической валидации системы искусственного интеллекта».

Следующей особенностью клинических испытаний СИИ можно считать необходимость *формирования набора данных* с определенными признаками для последующего тестирования данного ПО как медицинского изделия. Особую значимость в этом случае приобретает достоверность источника данных. Необходимо определиться с реестром хранилищ данных, пригодных для клинических испытаний. Набор данных для конкретного клинического испытания

формируется, как правило, с использованием официально утвержденных информационно-аналитических систем (например, ЕРИС ЕМИАС [7]).

Требования к набору данных для КИ могут быть сформулированы следующим образом:

1. Набор данных должен соответствовать области применения СИИ, описанной в технической и эксплуатационной документации на медицинское изделие.

2. Набор данных должен охватывать все возможные показания к применению исследуемого СИИ, указанные в технической и эксплуатационной документации.

3. Набор данных должен предоставлять возможность формировать результаты работы СИИ, наглядно представлять их с возможными дополнительными комментариями и пояснениями.

4. Набор данных должен содержать обезличенные данные и соответствовать этическим принципам и нормам в отношении пациентов, вовлеченных в процесс испытаний СИИ.

Набор данных только тогда становится пригодным для клинических испытаний, когда корректность этих данных подтверждена. Подтверждение корректности данных осуществляется специалистами-экспертами (врачами-экспертами). Отсюда вытекает следующая особенность клинических испытаний СИИ – необходимость *верификации* набора данных. В процессе верификации набора данных удаляются все данные, имеющие недостоверные признаки, ошибки в описании, не соответствующие требованиям технической и эксплуатационной документации конкретного СИИ как медицинского изделия. В качестве метода для верификации набора данных можно также применять метод «золотого стандарта» (ground truth) для целевой патологии, повторное исследование пациента через определенное время, результаты патогистологических, иммунологических исследований и др. (ГОСТ Р 59921.5-2022). Именно верифицированный набор данных подлежит использованию во всех дальнейших процедурах клинических испытаний конкретной СИИ.

При формировании набора данных для клинических испытаний СИИ осуществляется использование персональных данных субъектов, вовлеченных в процесс клинических испытаний. В связи с этим указанная деятельность должна проходить с соблюдением требований Федерального закона Российской Федерации от 27.07.2006 № 152-ФЗ «О персональных данных», регулирующего деятельность по обработке (использованию) персональных данных. В целях обеспечения защиты персональных данных сформированный набор данных должен быть деидентифицирован (обезличен) для цели дальнейшей обработки. Таким образом, возникает еще одна особенность клинических испытаний СИИ – необходимость деидентификации сформированного набора данных. Именно

эти данные, принадлежность которых конкретному субъекту персональных данных невозможно определить без дополнительной информации, подлежат дальнейшему использованию в процессе испытаний [5, 8].

Клиническая валидация завершается *расчетом характеристик (показателей) СИИ по результатам выполненной ей обработки верифицированного набора данных и сопоставления полученных результатов с данными, указанными в технической и эксплуатационной документации медицинского изделия* (см. раздел «Набор характеристик для оценки эффективности системы искусственного интеллекта»).

Важно отметить, что клиническая валидация ПО с ИИ проводится на верифицированном наборе данных, который не был использован ранее для обучения, настройки и первичной оценки СИИ.

Набор характеристик для оценки эффективности системы искусственного интеллекта

Общие определения

При оценке точности с использованием СИИ должен быть выполнен автоматизированный анализ (если применимо по назначению СИИ) набора данных (индекс-тест). Результаты разметки набора биомедицинских данных в соответствии с известной методологией – референс-тестом. Для сопоставления результатов индекс- и референс-теста составляют четырехпольную таблицу для целевой патологии и определяют абсолютное значение для каждого вида результатов (таблица 1). Далее производят выбор и расчет релевантных метрик с определением достоверности в рамках 95 % доверительного интервала.

Таблица 1 – Виды результатов применения ИИ в соответствии с четырехпольной таблицей (на примере ответа, представляющего собой бинарную переменную)

		Истинный показатель	
		Положительно	Отрицательно
Результат СИИ	Положительно	Истинно положительный (ИП) <i>СИИ выявила патологию при ее наличии</i>	Ложноположительный (ЛП) <i>СИИ выявила патологию при ее отсутствии</i>
	Отрицательно	Ложноотрицательный (ЛО) <i>СИИ не выявила патологию при ее наличии</i>	Истинно отрицательный (ИО) <i>СИИ не выявила патологию при ее отсутствии</i>

В зависимости от решаемой задачи могут применяться различные сочетания метрик. Например, анализ характеристической кривой и определение коэффициента согласованности классификаций и т.д. Такие основные показатели, как чувствительность, специфичность, площадь под характеристической кривой (ROC AUC), должны быть использованы для характеристики СИИ [9].

Для сопоставления эффективности индекс-теста относительно референс-теста применяется классический набор показателей, характеризующих диагностическую результативность (таблица 2).

Таблица 2 – Стандартный набор метрик диагностической ценности

Показатель	Значение	Формула
Чувствительность	Вероятность того, что индекс-тест окажется позитивным при наличии признака	$ИП / (ИП + ЛО)$
Специфичность	Вероятность того, что индекс-тест окажется негативным при отсутствии признака	$ИО / (ИО + ЛП)$
Точность (общая валидность)	Уровень соответствия полученных результатов индекс-теста действительным значениям референс-теста	$(ИП + ИО) / (ИП + ИО + ЛП + ЛО)$
Отношение правдоподобия положительного результата	Математически отображает, во сколько раз при положительном результате индекс-теста вероятность наличия целевой патологии превышает вероятность его отсутствия	$Чувствительность / (1 - специфичность)$
Отношение правдоподобия отрицательного результата	Математически отображает, во сколько раз при отрицательном результате индекс-теста вероятность наличия целевой патологии превышает вероятность ее отсутствия	$(1 - чувствительность) / специфичность$
Прогностическая ценность положительного результата	Вероятность заболевания при положительном результате индекс-теста	$ИП / (ИП + ЛП)$
Прогностическая ценность отрицательного результата	Вероятность отсутствия заболевания при отрицательном (нормальном) результате индекс-теста	$ИО / (ИО + ЛО)$
Частота ложных срабатываний	Доля неверных срабатываний к общему числу объектов за пределами класса	$1 - специфичность$
F-мера	Совместная оценка чувствительности и специфичности	$2 * специфичность * чувствительность / (специфичность + чувствительность)$

Характеристическая кривая (ROC)

Характеристическая кривая – отображение соотношения между долями объектов от общего количества носителей признака, верно классифицированными как несущие признак (то есть чувствительностью), и долями объектов от общего количества объектов, не несущих признака, ошибочно классифицированных как несущие признак (то есть специфичностью) при варьировании порога активации. Метрика диагностической ценности: площадь под кривой (AUC) – площадь, ограниченная ROC-кривой и абсциссой.

Классическая ROC-кривая: график зависимости чувствительности от 1 – специфичности (по оси абсцисс – 1 – специфичность, по оси ординат – чувствительность).

При анализе ROC-кривой необходимо воспользоваться пороговым значением, указанным в документации производителя, или определить оптимальные значения порога активации (так называемого cut-off). Переопределение порогового значения возможно тогда, когда это подтверждено производителем и отражено в итоговой документации клинических испытаний. В отдельных случаях переопределение порогового значения может привести к изменению качества работы СИИ, например, когда СИИ состоит из ансамбля алгоритмов. В случае, если производителем допускается переопределение пороговых значений, то для этого могут быть использованы различные методики. В качестве примера (но не обязательного требования) можно привести следующие:

1) минимальное расстояние от верхнего левого угла до ROC-кривой (минимум d);

2) индекс Юдена (Youden's index), который отображает максимальное расстояние от диагональной линии до ROC-кривой (рисунок 2).

Приоритет той или иной методике определения cut-off должен отдаваться в зависимости от целей испытания.

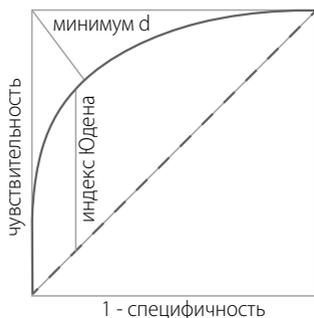


Рисунок 2 – Методика определения оптимальных значений порога активации (cut-off)

В анализе используют также разновидность ROC-кривой – precision-recall (PR)-кривую, которая строится как зависимость прогностической ценности положительного результата (precision) по оси ординат и чувствительности (recall) по оси абсцисс. Критерием качества СИИ при использовании данной метрики является площадь под PR-кривой (AUC-PR).

Коэффициент соответствия классификации

Согласованность (конкордантность) – математическое отображение согласованности классификации двух экспертов относительно одного явления. Показатель согласованности определяется по формуле Коэна (каппа Коэна) (1):

$$K = (P_0 - P_e) / (1 - P_e), \tag{1}$$

где P_0 – доля случаев, когда измерения совпадали, и P_e – ожидаемая доля случаев случайного совпадения.

Порядок расчета коэффициента соответствия классификации приведен в таблице 3.

Таблица 3 – Порядок расчета коэффициента соответствия классификации

		Эксперт № 1		
		1	2	Всего
Эксперт № 2	1	P_{11}	P_{12}	$P_{1\cdot}$
	2	P_{21}	P_{22}	$P_{2\cdot}$
	Всего	$P_{\cdot 1}$	$P_{\cdot 2}$	1

$$P_0 = P_{11} + P_{22} \tag{2}$$

$$P_e = P_{\cdot 1} P_{1\cdot} + P_{\cdot 2} P_{2\cdot}$$

где $P_{11}, P_{12}, P_{22}, P_{21}$ – доля случаев, когда мнения экспертов совпали; $P_{\cdot 1}, P_{\cdot 2}, P_{1\cdot}, P_{2\cdot}$ – доля случаев случайного совпадения.

Дополнительные метрики оценки для сегментации медицинских данных

Стандартная метрика кросс-энтропии – или логарифмическая функция потерь log-loss. Кросс-энтропия измеряет расхождение между двумя вероятностными распределениями. Если кросс-энтропия велика, это означает, что разница между двумя распределениями велика, а если кросс-энтропия мала, то распределения похожи друг на друга.

На рисунке 3 приведена схема обозначений для расчета метрик: на медицинских данных (I) выполнена эталонная сегментация экспертом (G), а также сегментация с использованием СИИ (M).

Дополнительно может быть использован ряд из 13 метрик: DICE, JAC, Kappa, Accuracy, Recall, Precision, Volumetric Similarity, Matthews Correlation Coefficient, Matthews Correlation Coefficient Normalized, Distance, Specificity, Balanced Accuracy, Adjusted Rand index. Выбор метрики будет зависеть от конкретной клинической задачи, а ее обоснование фиксируется в документации клинических испытаний. Например, для определения точности локализации патологических областей на рентгенограмме будут оптимальны метрики: Recall, Matthews Correlation Coefficient, Distance.

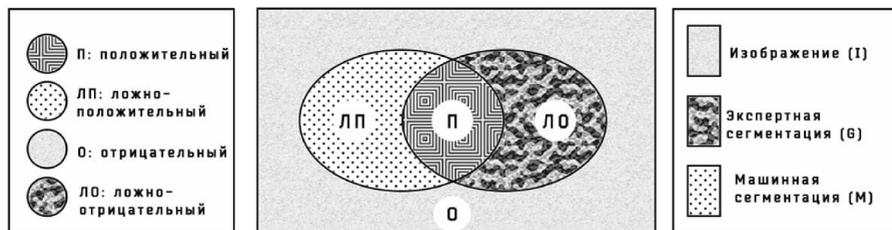


Рисунок 3 – Обозначение для расчета метрик сегментации медицинских данных

В таблице 4 приведен перечень метрик, которые применяют для оценки сегментации медицинских данных.

Таблица 4 – Перечень метрик для оценки сегментации медицинских данных

Метрики подобия	Разностные метрики
<p>Коэффициент подобия Дайса</p> $DSC = \frac{2 M \cap G }{ M + G }$	<p>Симметричный коэффициент разницы объема</p> $SVD = 1 - DSC$
<p>Коэффициент подобия Жаккара</p> $JSC = \frac{ M \cap G }{ M \cup G }$	<p>Ошибка наложения</p> $VOE = 1 - JSC$
<p>ИП доля объема</p> $TPVF(Rec) = \frac{ M \cap G }{ G }$	<p>ЛО доля объема</p> $FNVF = 1 - TPVF = \frac{ G \setminus M }{ G }$

Продолжение таблицы 4

Метрики подобия	Разностные метрики
ИО доля объема $TNVF = \frac{ I - M \cup G }{ I - G }$	ЛП доля объема $TNVF = 1 - TNVF = \frac{ M \setminus G }{ I - G }$
Точность $Prec = \frac{ M \cap G }{ M }$	-
-	Относительная разница объема $RVD = \frac{ M - G }{ G }$
-	Расстояние Хаусдорфа $HD = \max \left\{ \max_{x \in \partial G} d(x, \partial M), \max_{x \in \partial M} d(x, \partial G) \right\}$
-	$ASSD = \frac{\sum_{x \in \partial G} d(x, \partial M) + \sum_{x \in \partial M} d(x, \partial G)}{ \partial G + \partial M }$

ЗАКЛЮЧЕНИЕ

Клинические испытания – критически важный этап процесса жизненного цикла СИИ как медицинского изделия. Это одна из форм подтверждения соответствия медицинского изделия установленным требованиям. Такие испытания являются обязательными для успешной регистрации СИИ.

По завершении клинических испытаний в отчетной документации выносятся заключение о том, насколько эффективно и безопасно медицинское изделие, а также – насколько оправдано использование СИИ в дальнейшей медицинской практике.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Письмо Федеральной службы по надзору в сфере здравоохранения от 13.02.2020 № 02И-297/20 "О программном обеспечении".
2. Методические рекомендации по порядку проведения экспертизы качества, эффективности и безопасности медицинских изделий (в части программного обеспечения) для государственной регистрации в рамках национальной системы. М. : ФГБУ «ВНИИИИМТ» Росздравнадзора, 2021. 33 с.
3. Хельсинкская декларация Всемирной медицинской ассоциации «Этические принципы проведения медицинских исследований с участием человека в качестве субъекта» (принята на 18-й Генеральной Ассамблее Всемирной медицинской ассоциации). Финляндия, Хельсинки, июнь 1964 г.
4. Клинические испытания программного обеспечения на основе интеллектуальных технологий (лучевая диагностика) / сост. С. П. Морозов, А. В. Владзимирский, В. Г. Кляшторный [и др.] // Серия «Лучшие практики лучевой и инструментальной диагностики». М., 2019. Вып. 57. 51 с.
5. Павлов Н. А., Андрейченко А. Е., Владзимирский А. В. [и др.]. Эталонные медицинские датасеты (MosMedData) для независимой внешней оценки алгоритмов на основе искусственного интеллекта в диагностике // Digital diagnostics. 2021. Т. 2, № 1. С. 49–65.
6. 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. Vol. 13, № 3. P. 6–14.
7. Morozov S. With the unified radiological information system (URIS), the Moscow health care department has laid the foundations for connected care // *HealthManagement.org*. 2021. № 21(6). P. 348–351. URL: <https://healthmanagement.org/c/enterprise/issuearticle/with-the-uris-the-moscow-health-care-department-has-laid-the-foundations-for-connected-care> (дата обращения: 01.03.2023).
8. Andreychenko A., Logunova T., Gombolevskiy V. et al. A methodology for selection and quality control of the radiological computer vision deployment at the megalopolis scale // *medRxiv*. 2022. February 14. URL: <https://www.medrxiv.org/content/10.1101/2022.02.12.22270663v1> (дата обращения: 01.03.2023).
9. Зинченко В. В., Хоружая А. Н., Шарова Д. Е. [и др.]. Стандартизация в области регулирования технологий искусственного интеллекта в российском здравоохранении // *Казанский медицинский журнал*. 2021. Т. 102, № 6. С. 923–933.

Приложение А

РЕКОМЕНДОВАННАЯ ФОРМА ОТЧЕТА О ПРЕДВАРИТЕЛЬНЫХ КЛИНИКО-ТЕХНИЧЕСКИХ ИСПЫТАНИЯХ СИСТЕМЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Таблица А.1 – Рекомендованное содержание отчета о ПКТИ

Пункт	Пояснение
Медицинская организация	Полное название медицинской организации, на базе которой проведены ПКТИ
Контактная информация	Контактная информация медицинской организации (адрес, телефон, сайт, электронная почта)
Даты проведения ПКТИ	Диапазон дат проведения ПКТИ
Резюме	Структурированное изложение дизайна исследования, материалов, методов, результатов и выводов
Цель, задачи и конечные точки ПКТИ	Структурированное указание целей, задач выполнения ПКТИ и ожидаемого конечного результата
Набор данных	Детальная характеристика использованных для ПКТИ наборов данных
Вид данных	Вид(ы) данных (медицинская документация, результаты исследований и проч.), модальности для исследований, другие характеристики
Количество включенных клинических случаев	Количество включенных клинических случаев (пациентов, результатов исследований и др.)
Характеристика выборки	Характеристика популяции (расовые, гендерно-демографические, иные характеристики)
Характеристика набора данных и разметки	Где и когда сформирован набор данных с указанием ключевых характеристик (государственная регистрация базы данных, сведения о деперсонализации, о наличии информированного добровольного согласия пациентов, критерии включения/исключения, источники клинических случаев). Способ подготовки (разметки) набора данных
Характеристика патологии в наборе данных	Целевая патология и диагностические группы с распределением по ним. Способ верификации и наличие соответствующей информации в наборе данных
Способ формирования набора данных	Сформирован последовательно, случайно или другими способами. Обоснование размера выборки
Источники данных	Количество и локализация медицинских учреждений, послуживших источниками клинических случаев, включенных в ПКТИ

Продолжение таблицы А.1

Пункт	Пояснение
Отметка о независимости ПКТИ	Набор данных, использованный для ПКТИ, не должен как полностью, так и частично применяться для обучения или калибровки индекс-теста
Тестирование СИИ на наборе данных	Информация об использовании, инсталляции, организации доступа и прочие характеристики тестирования СИИ на наборе данных
Методы проведения ПКТИ	Краткое общее описание процесса исследования; возможно использование схем, диаграммы CONSORT
Таблица результатов	Комбинированная таблица результатов сравнения показателя набора данных и результата тестирования СИИ на наборе данных
Порог активации	Точка отсечения (cut-off) для тестирования СИИ на наборе данных, включая разделение исходов на «ожидаемые» и «непредвиденные»
Показатели эффективности	Метрики диагностической точности с 95 % доверительным интервалом (чувствительность, специфичность, общая точность, площадь под характеристической кривой и т.д.)
Ограничения	Любые ограничения ПКТИ, включая источники систематических ошибок, статистических неточностей и обобщений. Любые значимые различия в методиках формирования набора данных и тестирования СИИ на наборе данных
Выводы	Краткое обобщение результатов
Источники финансирования ПКТИ	Указание источника финансирования работ, связанных с проведением ПКТИ
Прочее	Произвольная дополнительная информация
Список исполнителей	Список сотрудников учреждения, проводивших ПКТИ (Ф. И. О., должность, ученое звание/степень)
Дата подписания отчета	–
Подпись ответственного исполнителя	Личная подпись ответственного исполнителя ПКТИ, Ф. И. О.
Подпись руководителя медицинской организации	Личная подпись руководителя учреждения, Ф. И. О.
Печать медицинской организации	–

ДЛЯ ЗАМЕТОК

ДЛЯ ЗАМЕТОК

Серия «Лучшие практики лучевой и инструментальной диагностики»

Выпуск 113

Составители:

*Васильев Юрий Александрович
Владзимирский Антон Вячеславович
Шарова Дарья Евгеньевна
Зинченко Виктория Валерьевна
Заюнчковский Сергей Юрьевич
Ахмад Екатерина Сергеевна
Гусев Александр Владимирович
Арзамасов Кирилл Михайлович*

**КЛИНИЧЕСКИЕ ИСПЫТАНИЯ
СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА
(ЛУЧЕВАЯ ДИАГНОСТИКА)**

Методические рекомендации

2-е издание, переработанное и дополненное

Отдел координации научной деятельности ГБУЗ «НПКЦ ДиТ ДЗМ»
Технический редактор А. И. Овчарова
Компьютерная верстка Е. Д. Бугаенко

ГБУЗ «НПКЦ ДиТ ДЗМ»
127051, г. Москва, ул. Петровка, д. 24, стр. 1



 +7 (495) 276-04-36

 npcmr@zdrav.mos.ru

 telemedai.ru