

Технологический конкурс в целях реализации Национальной
технологической инициативы

AI'M Doctor

по созданию интеллектуальной системы поддержки принятия врачебных
решений

Технический регламент (проект)

1.2

г. Москва
2023

Оглавление

1. Термины и определения	3
2. Общие положения	5
3. Условия участия в конкурсе	5
4. Структура Конкурса	6
5. Конкурсная задача и методика оценивания Продуктов Команд	7
6. Общие требования к процессам Испытаний на всех этапах	8
7. Отборочный этап	11
8. Полуфинальный этап	13
9. Финальный этап	15
10. Судейство Конкурса	17
11. Ограничения и дисквалификации	18
12. Обеспечение безопасности Конкурса	19
13. Дополнительные условия	20
Приложение 1. Интегральный критерий и метрики оценивания Совокупностей данных	22
Приложение 2. Типы сессий обмена с Платформой в рамках Отборочного этапа Конкурса	34

1. Термины и определения

- 1.1. В настоящем Техническом регламенте помимо терминов, перечисленных ниже, используются в том числе термины в значении, определенном в разделе «Общая информация» Конкурсного задания технологического конкурса в целях реализации Национальной технологической инициативы по созданию интеллектуальной системы поддержки принятия врачебных решений (далее – Конкурсное задание).
- 1.2. **Алгоритмическая квалификация** - процесс взаимодействия СППВР с Платформой на Испытаниях Отборочного этапа.
- 1.3. **Анонимизированные данные** – данные, из которых невозможно определить принадлежность данных конкретному физическому лицу - пациенту (субъекту персональных данных), сформированные на основе электронных медицинских карт пациентов путем уничтожения из них персональных данных и обрабатываемые в виде, исключающем возможности их деобезличивания.
- 1.4. **Валидационный датасет** - набор анонимизированных данных предназначенный для проведения Испытаний конкурса. Различают Валидационный датасет Отборочного, Полуфинального и Финального этапов, которые отличаются с точки зрения содержания и структуры. Валидационные датасеты состоят из **элементов, представляющих собой медицинские данные (Эпикризы)**. Валидационный датасет Финального этапа - состоит из Валидационного датасета Полуфинального этапа **и Эпикризов для финальных Испытаний**.
- 1.5. **Конкурс** – **совокупность** мероприятий, проводимых в целях реализации Национальной технологической инициативы по созданию интеллектуальной системы поддержки принятия врачебных решений AI'm Doctor. Предметом конкурса является выбор наилучшего решения для автоматической постановки заключительного клинического диагноза.
- 1.6. **Массовый технический сбой** в процессе проведения Испытаний фиксируется Технической комиссией в случае, если три и более элемента Валидационного датасета оказались просроченными (не были возвращены на Платформу в отведенное время) одновременно у всех Команд, участвующих в Испытаниях.
- 1.7. **МКБ 10** - Международная классификация болезней 10-го пересмотра

- 1.8. **Ознакомительный датасет** - набор анонимизированных данных (Эпикризов), предназначенный для демонстрации структуры и вида данных Валидационных датасетов (Эпикризов) Отборочного, Полуфинального и Финального этапов. Ознакомительный датасет может содержать все необходимые данные для постановки диагноза и сам диагноз (включая осложнения и сопутствующие заболевания). Структуры Ознакомительных и Валидационных датасетов не являются полностью идентичными. Оргкомитет имеет право изменять количество эпикризов Ознакомительного датасета.
- 1.9. **Платформа Конкурса (Платформа)** - распределенная информационная система, обеспечивающая доступ к данным, аналитической информации и сервисам для всех категорий пользователей, вовлечённых в участие в Конкурсе, а также для реализации процесса автоматического объективного оценивания Совокупностей данных в соответствии с математическим аппаратом Интегрального критерия.
- 1.10. **Программный комплекс участника (ПКУ, Продукт)** – разработанный Командой Участника Конкурса комплекс программного или программно-аппаратного обеспечения, который в автоматическом режиме принимает на входе эпикризы и выдаёт на выходе элементы заключительного клинического диагноза по установленной форме. В рамках Конкурса Программный комплекс участника и Система поддержки принятия врачебных решений (СППВР) являются синонимами.
- 1.11. **Сессия** - обмен информацией между СППВР Команды и Платформой в рамках решения задач на Отборочном, Полуфинальном и Финальном этапах. Запуск, остановка и выполнение действий регламентируется технической документацией в базе знаний Платформы.
- 1.12. **Совокупность данных** - множество данных, образующих комплексный результат деятельности СППВР в рамках какого-либо этапа Конкурса. Является закрытым элементом Платформы и не подлежит разглашению до момента окончания соответствующего этапа Конкурса.
- 1.13. **Система проверки решений (СПР)** - автоматическая система проверки, которая принимает множество данных о результатах работы СППВР и уровнях технологических барьеров; осуществляет расчет интегрального критерия и формирует аналитический отчет. СПР является частью Платформы.
- 1.14. **Тестовые испытания** - процесс, симулирующий отборочные, полуфинальные или финальные мероприятия и направленный на выявление различных ошибок, которые могут возникнуть как в тестируемых СППВР, так и в программно-техническом обеспечении упомянутых мероприятий.

- 1.15. **Судья** – член Судейской коллегии.
- 1.16. **Частичное преодоление Технологического барьера** – решение Конкурсной задачи Финального этапа Продуктом разработки при участии Клинического специалиста согласно условиям испытаний Финального этапа, определенным в п.5 Конкурсного задания и Техническом регламенте, с точностью не ниже порогового значения, определенного Экспертной комиссией, состоящей из профильных врачей высшей категории и экспертов по данным нозологиям, и требованиями описанными в Техническом регламенте.
- 1.17. **Эпикриз** - совокупность обезличенных (анонимизированных) или синтетических данных о пациенте, содержащая как результаты врачебных осмотров, так и результаты **лабораторных и инструментальных исследований**.

2. Общие положения

- 2.1. Настоящий Технический регламент определяет требования к процедуре проведения Конкурса, обеспечению и другим техническим параметрам Конкурса.
- 2.2. Настоящий Технический регламент является документом, детализирующим Конкурсное задание, опубликованное на официальном Сайте Конкурса, по методике и способу проведения Испытаний, а также процедуре определения результата и победителя Конкурса. Настоящий документ не противоречит Конкурсному заданию, но решает задачу детализации Конкурсного задания до технического уровня. Технический регламент определяет требования к процедуре проведения Испытаний Конкурса, Программным комплексам участников, процедуре контроля и алгоритму определения качественного уровня разработок Команд.
- 2.3. Основная цель настоящего документа: обеспечение единого набора документации и регламента участия в Конкурсе, обеспечивающих прозрачность и объективность как для Участников, так и для сторонних наблюдателей.
- 2.4. **Настоящий Технический регламент может быть изменен решением Организационного комитета. Изменения вносятся не позднее 30 календарных дней до мероприятия, затрагиваемого изменениями. Любые изменения, вносимые в Технический регламент, публикуются на Сайте Конкурса.**

3. Условия участия в конкурсе

- 3.1. Требования к Участникам и процессу формирования Команд описаны в п.4 Конкурсного задания Конкурса

- 3.2. Для участия в конкурсе необходимо подать заявку путем заполнения электронной формы на Сайте конкурса. Данный процесс подробно изложен в п.3.2 Конкурсного задания Конкурса
- 3.3. Подавая заявку на участие в Конкурсе, Участник соглашается с требованиями Конкурсного задания и Технического регламента Конкурса и обязуется следовать **указанным в них условиям** и требованиям.
- 3.4. Все Участники, на всех этапах Конкурса, а также при разработке своих **Продуктов** и получаемых с помощью **Продуктов Совокупностей данных** обязаны руководствоваться целями и задачами Конкурса, положениями Конкурсного задания и Технического регламента.
- 3.5. Дополнительные условия участия описаны в п.3.4 Конкурсного задания Конкурса.

4. Структура Конкурса

- 4.1. Конкурс проводится в три этапа:
 - 4.1.1. Отборочный этап. Формирование Команд и подтверждение Участниками уровня подготовки, минимально-необходимого для участия в последующих этапах. Отборочный этап подробно описан в п.7 настоящего Технического регламента, а также в п.5.2 Конкурсного задания Конкурса.
 - 4.1.2. Полуфинальный этап. Проверка способности СППВР Команд выполнять задачу Конкурса, демонстрируя повторяемый качественный результат, а также отбор Команд для Финального этапа. Полуфинальный этап подробно описан в п.8 настоящего Технического регламента, а также в п.5.3 Конкурсного задания Конкурса.
 - 4.1.3. Финальный этап. Проверка способности СППВР Команд выполнять задачу Конкурса в реальных условиях применения современных интеллектуальных медицинских систем. Финальный этап подробно описан в п.9 настоящего Технического регламента, а также в п.5.4 Конкурсного задания Конкурса.
- 4.2. **В рамках каждого этапа Конкурса проводятся Испытания.** Испытания могут иметь разный формат (быть очными или заочными; проводиться дистанционно или оффлайн). Под дистанционными Испытаниями понимается процесс, когда все Команды подключаются одновременно к Платформе и работают в параллельном режиме над решением определенной etapом задачи. **Ход дистанционных Испытаний на**

Полуфинальном этапе и Испытаний в режиме оффлайн на Финальном этапе контролируется Судейской коллегией и Технической комиссией.

5. Конкурсная задача и методика оценивания **Продуктов Команд**

- 5.1. Обоснование темы Конкурса и границ конкурсной задачи приведены в п.2. Конкурсного задания Конкурса.
- 5.2. Цель проведения Конкурса (согласно п.1.4 КЗ) – преодоление технологического барьера в области искусственного интеллекта (далее - ИИ) в медицине.
- 5.3. Технологическим барьером для Конкурса (согласно п.3.1.3 КЗ) является разработка технологического решения на основе ИИ в виде комплекса программного обеспечения, способного осуществить постановку заключительного клинического диагноза с приемлемой для целей здравоохранения точностью, достоверностью и эффективностью на основе анализа комплекса клиничко-лабораторно-диагностических и анамнестических данных и медицинских знаний и клинических рекомендаций.
- 5.4. Критерий преодоления Технологического барьера (согласно 5.4.20 КЗ) - процент точной постановки диагноза по основному заболеванию на всех датасетах в процессе решения задачи Финального этапа не ниже порогового значения, определенного Экспертной комиссией и утвержденного Техническим регламентом, что является необходимым условием для победы в Финальном этапе.
- 5.5. В качестве пороговых значений Технологического барьера, учитываемых Интегральным критерием, используется чувствительность и специфичность в рамках конкретной нозологии с определенным в Интегральном критерии уровнем доверия.

Нозология	Рак легкого	Туберкулез	Бактериальная и вирусная пневмония (вкл. SARS-CoV-2)	ХТЭЛГ	ХОБЛ
Чувствительность	78,000%	79,000%	75,000%	81,000%	75,000%
Специфичность	89,000%	88,000%	79,000%	93,000%	84,000%

- 5.6. В случае, если Продукт разработки после прохождения финальных Испытаний не соответствует требованиям п 5.5 настоящего Технического регламента – **возможно только Частичное преодоление Технологического барьера.** Условия присуждения Частичного преодоления Технологического барьера описаны в Приложении 1.

- 5.7. Для качественного ранжирования результатов Команд в рамках Конкурса и определения факта преодоления Технологического барьера используется математический аппарат, обеспечивающий комплексное многофакторное оценивание Продуктов - Интегральный критерий. Подробное описание метрик и методик оценивания приведено в Приложении 1 настоящего Технического регламента.

6. Общие требования к процессам Испытаний на всех этапах

- 6.1. Для обеспечения работы на всех этапах Конкурса используется распределенная информационная система Конкурса - Платформа. Все процессы испытаний на всех этапах конкурса подразумевают взаимодействие Продуктов с Платформой Конкурса. Участие Команды во всех процессах Конкурса возможно и разрешено только с использованием документированного функционала и инструментария Платформы.
- 6.2. Команда получает доступ к Платформе после успешного прохождения процедуры регистрации в Конкурсе. В рамках Платформы Командам доступны:
- 6.2.1. Личный кабинет с необходимыми исходными данными, **включая базу знаний Платформы**, и результатами прохождения этапов.
 - 6.2.2. Датасеты конкурса. Перечень и наполнение датасетов зависит от текущего этапа Конкурса.
 - 6.2.3. Программный модуль для обмена данными с Платформой (включая полный исходный код), а также, подробное описание процесса взаимодействия с Платформой. Информация предоставлена на Github по следующей ссылке: <https://github.com/ai-med-tools/baseline> (ссылка может быть обновлена. В данном случае, актуальная ссылка размещается на Сайте).
- 6.3. Процессы Испытаний подразумевают общую итерационную последовательность действий **на Платформе**:
- Шаг 1. Платформа предоставляет набор данных для чтения **Продуктами Команд**;
 - Шаг 2. **Продукты Команд** анализируют полученные данные и формируют результат анализа в соответствии с требованиями этапа;
 - Шаг 3. Результат анализа (в виде файла утвержденного формата) передается **Продуктами Команд** на Платформу;

Шаг 4. Возврат на шаг 1.

- 6.4. В процессе Испытаний элементы Валидационного датасета (эпикризы в виде XML-файлов с входными анонимизированными данными пациентов, **которые могут содержать** жалобы пациента, анамнестические данные, протоколы исследований и консультаций с врачами) поступают на вход СППВР Команд.
- 6.5. **Коды основных заболеваний, разрешенные в рамках Конкурса:**

Нозология	МКБ 10
Туберкулез	A15 A16 A19
Рак легкого	C34
Бактериальная пневмония	J13 J14 J15 J18
Вирусная пневмония (включая SARS-Cov2)	J10 J12 U07 U07
Хроническая обструктивная болезнь легких (ХОБЛ)	J44
Хроническая тромбоэмболическая легочная гипертензия (ХТЭЛГ)	I27 I50 I50 I50 I26

- 6.6. Для участия в Полуфинальных и Финальных испытаниях Команда должна запустить Сессию обмена данными с Платформой до момента старта испытаний **и по команде Оргкомитета**. В случае подключения к Платформе после начала Испытаний, метрика для пропущенных Командой Эпикризов аннулируется, диагноз считается поставленным неверно.
- 6.7. **Временные интервалы, отводимые Командам на выполнение заданий в рамках Испытаний этапов, отсчитываются от момента публикации данных Платформой. При этом время ответа Команды определяется по факту получения данных Платформой. Командам стоит учитывать возможные сетевые задержки, не зависящие от Платформы, но влияющие на время получения ответа Платформой. Оргкомитет не несет**

ответственности за сбои в работе, возникшие по вине Интернет-провайдера на стороне Команды.

- 6.8. В рамках Полуфинального и Финального этапов проводятся Тестовые испытания. Тестовые испытания не являются обязательными для Команд и проводятся в том же порядке, что и Испытания, но на меньшей выборке данных. Результаты Продуктов на Тестовых испытаниях не оцениваются. Даты проведения Тестовых испытаний сообщаются Командам не позднее чем за 5 (пять) рабочих дней до момента проведения.
- 6.9. Судейская коллегия, Техническая комиссия и Экспертная комиссия проводит предварительную ревизию Совокупностей данных на предмет соответствия целям и задачам Конкурса. Результаты предварительной ревизии передаются в Оргкомитет. В случае, если Совокупность данных какой-либо из Команд не соответствует целям и задачам Конкурса - эта Совокупность данных аннулируется Оргкомитетом и расчет метрик Интегрального критерия по аннулированной Совокупности данных не производится.
- 6.10. В течение суток с момента окончания сессии Испытаний в личном кабинете Участника формируется отчет, отражающий предварительные результаты Команды. Итоговыми результатами считаются результаты, утвержденные Жюри Конкурса.
- 6.11. На протяжении всех этапов Конкурса Оргкомитетом предоставляется канал для оперативной коммуникации между Командами и организаторами.
 - 6.11.1. В качестве канала коммуникации могут использоваться инструменты мессенджеров или сред совместной работы.
 - 6.11.2. Члены Команд обязаны иметь в канале коммуникации имя пользователя (ник), соответствующий следующей структуре: **Название Команды_ФИО члена Команды полностью.**
- 6.12. Испытания Полуфинального и Финального этапов могут быть перезапущены в случае Массового технического сбоя по решению Оргкомитета, на основании соответствующих протоколов Судейской Коллегии и Технической Комиссии. В случае перезапуска:
 - 6.12.1. В качестве даты новых Испытаний может быть назначен другой день (но не позднее чем 7 дней с момента отмененного этапа), о чём Оргкомитет уведомляет отдельно.
 - 6.12.2. Используется новый Валидационный датасет, размер и наполнение которого определяется Оргкомитетом. Все предыдущие результаты Команд в рамках данного этапа могут быть аннулированы. Регламент действий в случае перезапуска определяется Оргкомитетом.

- 6.13. Использование общих программных модулей и установленных разработчиком Платформы способов взаимодействия с Платформой, является обязательным, поскольку обеспечивает равные условия для всех Участников Конкурса.
- 6.14. Участники Конкурса не имеют права самостоятельно вносить изменения в общие программные модули **Платформы**.

7. Отборочный этап

- 7.1. Данный раздел дополняет п.5.2 Конкурсного задания Конкурса
- 7.2. Задачей Отборочного этапа является формирование Команд, готовых технически и организационно к процессу работы в условиях Зачетных испытаний последующих этапов Конкурса.
- 7.3. Отборочный этап начинается с момента определенного Оргкомитетом и завершается подведением итогов. Испытания Отборочного этапа проводятся в заочном формате.
- 7.4. После одобрения Оргкомитетом заявки Команды на участие в Конкурсе, Команда получает доступ в Личный кабинет Участника на Платформе.
- 7.5. Команда должна загрузить с Платформы необходимые программные модули и развернуть клиентскую часть ПКУ для подключения к Платформе собственных СППВР.
- 7.6. В рамках Отборочного этапа Командам становится доступен Ознакомительный датасет. **В рамках Испытаний Отборочного этапа становится доступен Валидационный датасет.**
- 7.7. **В рамках Испытаний отборочного этапа Командам доступно два вида Сессий обмена с Платформой: тренировочные и алгоритмические. Описание форматов ответа СППВР в рамках Сессий указано в Приложении 2.**
- 7.8. **Для успешного прохождения Испытаний Отборочного этапа СППВР Команды должна пройти Алгоритмическую квалификацию:**
 - 7.8.1. **Запустить Сессию на Платформе с типом Сессии - **algorithmic** (алгоритмическая).**
 - 7.8.2. **Получить от Платформы 1 случайный элемент Валидационного датасета.**
 - 7.8.3. **Обработать элемент Валидационного датасета и вернуть на Платформу код МКБ 10 основного заболевания (см п.7.9).**

- 7.9. Задание Отборочного этапа в своей структуре уже содержит сведения о диагнозе (основном заболевании). Таким образом, задача Команды заключается в определении этих данных внутри структуры Эпикриза, считывании данных и возврате данных на Платформу.
- 7.10. Требования к ответу СППВР в рамках прохождения Алгоритмической квалификации:
- 7.10.1. Код МКБ 10 должен быть передан без пробелов и служебных символов.
- 7.10.2. Код МКБ 10 должен быть передан в строгом соответствии с Эпикризом (в том же регистре). Сравнение производится по базовой части кода без уточнения (до символа “.”).
- 7.11. Задание отборочного этапа может являться как образцом реального анонимизированного Эпикриза, так и набором синтетических данных, имеющих структуру реального Эпикриза.
- 7.12. Для проведения Алгоритмической квалификации Команда может открыть Сессию обмена данными с Платформой в любой момент в рамках периода Испытаний Отборочного этапа. При открытии Сессии сервер Платформы отправляет ПКУ задание и ожидает ответ от ПКУ в течение 60 секунд. Если Платформой не получен ответ от ПКУ в течение 60 секунд - Сессия закрывается.
- 7.13. Команда может проходить алгоритмическую квалификацию многократно в пределах времени, предоставляемого Оргкомитетом для этого процесса. Платформой накладываются ограничения на максимальную периодичность отправок в рамках Сессий (10 отправок в день).
- 7.14. Команда, успешно прошедшая Алгоритмическую квалификацию, считается прошедшей Испытания Отборочного этапа после утверждения результата Команды Оргкомитетом.

8. Полуфинальный этап

- 8.1. Данный раздел дополняет п.5.3 Конкурсного задания Конкурса.
- 8.2. Полуфинальный этап начинается с момента подведения итогов Отборочного этапа и завершается Испытаниями в дистанционном формате.
- 8.3. В процессе Полуфинального этапа Команде передается Ознакомительный датасет. Команда вправе использовать Ознакомительный датасет как обучающую выборку, однако, организатор Конкурса не гарантирует репрезентативность и достаточность выборки.

- 8.4. В рамках данного этапа задачей СППВР Команды является:
 - 8.4.1. Постановка предварительного клинического диагноза по каждому из Эпикризов Валидационной выборки на основе неполных данных.
 - 8.4.2. Постановка Заключительного клинического диагноза по каждому из Эпикризов Валидационной выборки.
- 8.5. Результаты работы СППВР, представленные в виде кодов МКБ 10 основного диагноза, выявленных осложнений и сопутствующих заболеваний для каждого пациента, и передаются на Платформу (в Систему проверки решений).
- 8.6. Система проверки решений сравнивает результат работы СППВР Команд с эталонным результатом, предварительно сформированным экспертной группой врачей.
- 8.7. По результатам сравнения рассчитываются метрики:
 - 8.7.1. точность (чувствительность и специфичность) выявления каждой из нозологий основного диагноза;
 - 8.7.2. точность выявления осложнений и сопутствующих заболеваний.
 - 8.7.3. **ресурсоемкости** постановки диагноза (для последующего **использования** на Финальном этапе Конкурса).
- 8.8. Испытания Полуфинального этапа проводятся на основе Валидационного датасета, имеющего объем **не менее 500** Эпикризов. Диагнозы в Валидационном датасете Полуфинального этапа отсутствуют.
- 8.9. **Каждый элемент Валидационного датасета представлен в двух версиях - неполные данные и полные данные. Состав версии с неполными данными может различаться в разных элементах датасета.**
- 8.10. Во время Полуфинальных испытаний элементы датасета становятся доступны (публикуются) по расписанию, начиная с момента старта, через равные промежутки времени в 30 секунд для обеспечения равных условий всех Участников Конкурса и исключения любых возможностей для организации ручной постановки диагноза. Расписание полуфинальных Испытаний объявляется не позднее, чем за 5 (пять) рабочих дней до их начала.
- 8.11. Полуфинальные испытания проходят в два подэтапа: предварительная постановка диагноза и окончательная постановка диагноза. Процесс реализуется следующим образом:
 - 8.11.1. Платформа публикует случайный **Эпикриз** в версии неполных данных.
 - 8.11.2. СППВР Команды анализирует **Эпикриз**, и на основе информации в нем возвращает предполагаемый основной диагноз в виде кода МКБ 10. На данную операцию отводится 30 секунд.
 - 8.11.3. Платформа публикует полную версию **текущего элемента Валидационного датасета.**

- 8.11.4. СППВР Команд анализирует файл, и на основе информации в нем возвращает:
- либо код МКБ 10 Основного заболевания (одна из шести утвержденных нозологий). Коды МКБ 10 Сопутствующих заболеваний и Осложнений;
 - либо значение “Не болен”, если пациент не болен ни одной из нозологий Конкурса.
- На данную операцию отводится 30 секунд.
- 8.12. В ходе Сессии обмена данными ПКУ должен оперативно загружать и обрабатывать файлы Эпикризов с сервера Платформы, затем отправлять файлы с диагнозами и дополнительной информацией в Систему проверки решений (СПР). При превышении порогового времени задержки в 30 секунд (после публикации данного файла) файл аннулируется для данной Команды, **диагноз считается поставленным неверно**.
- 8.13. Если аннулирование файла Эпикризов происходит у Команд по причине **Массового технического сбоя** на стороне Платформы, то Оргкомитет имеет право заменить эти файлы резервными, продлив Сессию обмена данными, но не более, чем на 100 (сто) файлов. В случае замены аннулированные файлы исключаются из расчета метрик Интегрального критерия для всех Команд.
- 8.14. По окончании циклов отправки/приема всего объема выборки Полуфинального датасета Сессии обмена данными между ПКУ и Платформой закрываются, все данные, полученные от Команды, сохраняются в Базе данных диагнозов, формируя закрытый датасет с полуфинальными **Совокупностями данных Команд**. С этого момента стартует процесс оценивания **результатов Команд**.
- 8.15. После предварительной ревизии **Совокупностей данных** запускается процесс расчета Интегрального критерия Команд.
- 8.16. По итогам Полуфинального этапа строится лидерборд — список Команд, ранжированный по убыванию Интегрального критерия (см. Приложение 1). Аннулированные решения не участвуют в финальном лидерборде.
- 8.17. К финальному этапу допускается не более 4 Команд. Для участия в финальном этапе Команда должна продемонстрировать решение превышающее пороговую точность, а также иметь один из 4 лучших значений Интегрального критерия.
- 8.18. Преодоление технологического барьера на Полуфинальном этапе Конкурса не является условием победы Команды в Конкурсе и определяет только возможность к переходу на Финальный этап.

- 8.19. Если по результатам Полуфинального этапа не может быть сформирован пул из 4 Команд для Финального этапа, то Полуфинальный этап может быть проведен повторно по решению Оргкомитета.

9. Финальный этап

- 9.1. Данный раздел дополняет п.5.4 Конкурсного задания Конкурса
- 9.2. В рамках данного этапа **Продуктам Команд** необходимо:
- 9.2.1. Определить заключительный клинический диагноз по всем Эпикризам валидационного датасета.
- 9.2.2. Иметь возможность работать в паре с Клиническим специалистом в режиме пошаговой постановки диагноза. На каждом шаге СППВР запрашивает необходимые данные Эпикриза и обосновывает Клиническому специалисту свое решение.
- 9.3. Критерии оценки СППВР включают в себя: точность постановки диагноза; **ресурсоемкость** постановки диагноза.
- 9.4. Финальный этап начинается с момента подведения итогов Полуфинального этапа и завершается очными Испытаниями с личным присутствием **членов Команды** на площадке Испытаний.
- 9.5. Финальные испытания проводятся на основе выборки Финального датасета, имеющего объем не менее **64 Эпикризов**. Диагнозы в Валидационном датасете финального этапа отсутствуют.
- 9.6. Концепция Финального этапа заключается в итерационной постановке диагноза в условиях неполных или недостаточных анонимизированных данных о пациенте. Изначально Команда получает первичный набор анонимизированных данных о пациенте, после чего должна запрашивать через Платформу необходимые данные, аргументируя свои решения Клиническому специалисту.
- 9.7. Финальный этап представляет собой процесс последовательной обработки элементов Финального датасета (Эпикризов пациентов) с целью постановки основного клинического диагноза, определения осложнений и сопутствующих заболеваний.
- 9.8. Обработка каждого элемента Финального датасета выполняется в 2 шага:
- 9.8.1. Шаг А: Итерационная постановка диагнозов в режиме ограниченного времени.
- 9.8.1.1. СППВР получает первичные анонимизированные данные о пациенте и анализирует данные. Набор и вид первичных

анонимизированных данных определяется после завершения Полуфинального этапа Конкурса

- 9.8.1.2. Если постановка диагноза на основе имеющейся информации невозможна, СППВР формирует комментарий Клиническому специалисту, содержащий основания для запроса дополнительной информации о пациенте (анализ, исследование и т.д.). Комментарий клиническому специалисту сохраняется на Платформе.
 - 9.8.1.3. СППВР запрашивает дополнительную информацию о пациенте. За одну итерацию можно запросить результаты только одного анализа или исследования.
 - 9.8.1.4. В случае если запрошенная информация присутствует в Эпикризе - она будет передана СППВР. В случае отсутствия запрошенной информации в Эпикризе СППВР необходимо выбрать иную траекторию постановки диагноза и перейти к п.9.8.1.2
 - 9.8.1.5. Шаги 9.8.1.1 - 9.8.1.3 повторяются:
 - до момента постановки основного диагноза, определения сопутствующих заболеваний и осложнений;
 - до момента остановки процесса со стороны СППВР ввиду невозможности поставить диагноз;
 - до момента истечения времени (1 минута).
 - 9.8.1.6. На Платформе сохраняется информация о траектории постановки диагнозов и комментарии Клиническому специалисту.
- 9.8.2. Шаг Б: Работа в паре с Клиническим специалистом
- 9.8.2.1. Определение пар СППВР-Клинический специалист происходит по методике п.5.4.17 Конкурсного задания Конкурса.
 - 9.8.2.2. Взаимодействие СППВР и Клинического специалиста осуществляется через Платформу.
 - 9.8.2.3. Платформа воспроизводит работу СППВР по траектории, сохраненной на шаге А.
 - 9.8.2.4. На каждой итерации Клинический специалист анализирует комментарий СППВР и адекватность запроса текущей ситуации. Задача клинического специалиста - проанализировать комментарии СППВР. На каждой итерации Клинический специалист оценивает запрос СППВР, исходя из двух возможных состояний: “корректно” и “некорректно”.

- 9.9. После окончания Испытаний решения Команд и Клинических специалистов сохраняются на Платформе, формируя закрытый датасет результатов Финального этапа.
- 9.10. Платформа производит расчет метрик интегрального критерия на выборке данных закрытого датасета результатов Финального этапа.
- 9.11. Испытания финального этапа проводятся очно на территории Организатора. В процессе проведения Финальных испытаний на площадке испытаний должны присутствовать не менее двух представителей Команды.
- 9.12. Процесс Финальных испытаний регламентируется **правилами внутреннего распорядка**, зависящими от условий Конкурсной площадки. Данные правила формируются Оргкомитетом Конкурса и доводятся до сведения Участников не позднее чем за 10 рабочих дней до начала Испытаний.

10. Судейство Конкурса

- 10.1. Судейская коллегия проверяет автоматически формируемые заключения о прохождении Командой алгоритмической квалификации и готовности к Испытаниям Полуфинального и Финального этапов.
- 10.2. Результаты Испытаний Полуфинального и Финального этапов верифицируются Технической комиссией и передаются на утверждение Оргкомитету.
- 10.3. Техническая комиссия проверяет корректность полученных от Команды данных, выявляет условия и причины аннулирований на стороне Платформы, если таковые имеются.
- 10.4. По окончании Финальных испытаний и процедур расчета интегрального критерия Судейская коллегия проверяет автоматически формируемые отчеты о **результатах Команд** и условия преодоления технологического барьера.
- 10.5. Процесс определения победителей подробно описан в п.6 и п.7 Конкурсного задания Конкурса.
- 10.6. Оргкомитет выносит итоговые решения по результатам Полуфинального и Финального этапов, которые передаются на рассмотрение и утверждение Жюри.

11. Ограничения и дисквалификации

- 11.1. Участники и **члены Команд** Конкурса обязаны следовать требованиям Конкурсного Задания и Технического регламента Конкурса
- 11.2. В процессе Конкурса Участники и **члены Команд** обязаны следовать принципам честности и нормам морали принятым в обществе при взаимодействии с организаторами или другими Участниками и **членами**

- Команд**, соблюдать требования этики, быть внимательным и доброжелательным в общении с другими Участниками, **членами Команд** и Оргкомитетом, поддерживать имидж организатора, воздерживаться от некорректных высказываний в отношении профессионализма других Участников, воздерживаться от оскорблений в адрес других Участников, Оргкомитета и партнеров Конкурса.
- 11.3. Участнику **и Командам** запрещается осуществлять экспертную разметку Тестового или Валидационного датасета вручную, с помощью краудсорсинга или любыми иными способами с привлечением труда людей.
 - 11.4. Запрещаются действия направленные на попытку определить персональные данные субъекта Эпикриза (пациента) путем анализа сходства результатов клинических исследований или любым другим путем. Подобные действия будут рассчитаны как нарушение Участником закона о персональных данных.
 - 11.5. Запрещаются действия направленные на попытку получения диагноза по какому-либо из Эпикризов каким-либо образом кроме способов явно разрешенных Техническим регламентом и Конкурсным заданием (получение диагнозов от третьих лиц; из сторонних баз данных и т.д.).
 - 11.6. Запрещается вносить изменения в общие программные модули Платформы или любым образом прямо или косвенно влиять на возможность подобного изменения. Исключением является только прямое обращение Участника в Оргкомитет с аргументированным предложением изменений в программных модулях.
 - 11.7. На всех этапах конкурса при работе с Платформой запрещаются любые обращения к Платформе с частотой выше 10 раз в секунду.
 - 11.8. В случае нахождения особенностей в математическом или алгоритмическом обеспечении Платформы, способных повлиять на результаты Конкурса, Команде запрещается использовать данные особенности в целях дискредитации Конкурса или получения выгоды.
 - 11.9. Запрещается использовать логические, алгоритмические и иные решения, противоречащие целям и задачам Конкурса. В случае подозрения в использовании подобных решений Оргкомитет имеет право дисквалифицировать Команду. Настоящее ограничение распространяется и на особенности **Продукта**, противоречащие целям и задачам Конкурса без умысла Команды. В случае сомнений соответствии решений целям конкурса Команда имеет право запросить официальную позицию Оргкомитета.
 - 11.10. В случае нарушения любого из перечисленных пунктов Оргкомитет по запросу Судейской коллегии или Технической комиссии имеет право отстранить Команду от участия во всех последующих этапах Конкурса, а

также дисквалифицировать Команду (аннулировать все полученные ранее результаты Команды и полностью прекратить участие Команды в Конкурсе).

12. Обеспечение безопасности Конкурса

- 12.1. Если в процессе Испытаний Команда отправляет на Платформу несколько ответов на один запрос, то алгоритмы Платформы учитывают только последний вариант.
- 12.2. В случае опоздания Команды на Испытания, Оргкомитет может принять решение о допуске к Испытаниям с текущего момента. Однако, в данном случае все отработанные данные Валидационного датасета (Эпикризы полученные другими Командами) считаются аннулированными для опоздавшей Команды. Аннулирование равнозначно постановке неверного диагноза для конкретного блока данных (конкретного Эпикриза).
- 12.3. Требования к безопасности и экологии, ограничения на раскрытие и распространение информации, права на интеллектуальную собственность и порядок изменения условий Конкурса описаны в п. 8 конкурсного задания Конкурса.
- 12.4. Деятельность Участников в рамках Конкурса должна соответствовать действующим на территории Российской Федерации природоохранным нормам, требованиям техники безопасности.
- 12.5. На финальном этапе Конкурса Оргкомитет может предоставить Командам правила по технике безопасности и охране окружающей среды, с которыми все члены Команд должны ознакомиться под подпись и соблюдать.

13. Дополнительные условия

- 13.1. Оргкомитет оставляет за собой право обновлять Ознакомительный и Валидационный датасеты, а также исходные коды общих программных модулей, но не позднее, чем за 10 (десять) календарных дней до Финальных испытаний. В случае обновления Оргкомитет оповещает всех Участников путем публикации на сайте Платформы, сообщая цель и суть произведенных модификаций.
- 13.2. Оргкомитет имеет право приостановить ход Зачетных испытаний любого из этапов, например, для выяснения причины технического сбоя на стороне Платформы и его оперативного устранения. Всем Командам приходит сообщение «Пауза», после чего сервер Платформы временно перестает выполнять штатные действия по реализации испытаний. По окончании паузы всем Командам приходит сообщение «Конец паузы», после чего сервер не более чем через 1 минуту возобновляет штатную работу. Суммарная продолжительность пауз во время Финальных

испытаний не может превышать 2 часов. В случае необходимости большей продолжительности Оргкомитет принимает решение о переносе Испытаний

- 13.3. Оргкомитет имеет право проверить решение Команды очно или провести дополнительные процедуры проверки, в случае если у Судейской коллегии или Технической комиссии возникнут основания считать Решение какой-либо Команды не соответствующим целям и задачам Конкурса
- 13.4. Оргкомитет имеет право публиковать результаты Решений Команд (включая траектории постановки диагнозов и интерпретации Клиническому специалисту) после испытаний Финального этапа на Сайте конкурса с целью обеспечения прозрачности в определении победителей и призёров.
- 13.5. Все технические инструкции по работе Платформы, а также параметры и способы работы Платформы являются предметом отдельных документов, которые публикуются на официальном сайте Конкурсов, или на ресурсах Платформы или ссылки на них ведут с Сайта или Платформы. Указанные документы **не противоречат Конкурсному заданию и Техническому регламенту**, но описывают технические параметры функционала Платформы и одинаково доступны всем Командам, в том числе для соблюдения равных условий участия Команд.
- 13.6. В соответствии с п. 8.2. КЗ данные, сформированные ПКУ во время Испытаний (но не само программное обеспечение, являющееся интеллектуальной собственностью Команд) передаются Организаторам на условиях открытой лицензии Creative Commons Attribution 4.0 International (CC BY 4.0) для последующей публикации на Сайте конкурса. В целях выполнения настоящего пункта, Участник гарантирует, что им получены согласия членов Команды на такую публикацию результатов. По запросу Организаторов Участник в течение 3 рабочих дней предоставляет письменные доказательства выполнения указанного условия.

Приложение 1. Интегральный критерий и метрики оценивания Совокупностей данных

1. Показатели и метрики используемые для оценки качества медицинской диагностики

В клиническом диагностическом тесте с бинарными исходами для каждого набора данных, соответствующих отдельному пациенту или субъекту исследований, ставится положительный или отрицательный диагноз, и этот диагноз сравнивается с известным истинным состоянием.

При постановке диагноза возможны четыре результата:

A - истинно положительный (пациент 'болен' и система диагностировала его как 'болен');

B - ложноположительный (пациент 'здоров', но система его диагностировала как 'болен');

D - истинно отрицательный (пациент 'здоров' и диагностирован системой как 'здоров');

C – ложноотрицательный (пациент 'болен', но диагностирован системой как 'здоров').

Таким образом если исследуется *N* пациентов, то *A* и *D* – количества верных, а *B* и *C* – количества неверных диагностических ответов, для которых справедливо:

$$N = A + D + B + C \quad (1)$$

Результаты диагностики (правильные и ошибочные) представляют матрицей ошибок (Рис.1) в которой *A*, *B*, *C* и *D* — количественные результаты диагностики системой по выборке пациентов *N*, соответствующих каждому результату диагностической классификации.

Нозология		<i>i</i>		<i>i</i> =[1,,6]
Результаты теста по нозологии <i>i</i>		СППВР		
		болен	не болен	всего
Факт	болен	TP 68 (A)	FN 7 (C)	75 (n1=A+C)
	не болен	FP 4 (B)	TN 41 (D)	45 (n0=B+D)
	всего	72	48	120 (N)

Рис 1. Матрица ошибок диагностического теста (бинарная классификация)

В медицинской практике показателями точности диагностических тестов являются чувствительность (истинно положительные результаты) и специфичность (истинные отрицательные результаты). Чувствительность диагностического теста — доля тех, у кого есть состояние, для которого диагностический тест верно определил как положительный, а специфичность диагностического теста — это доля тех, у кого нет состояния, для которого диагностический тест является отрицательным. Используя приведенную выше классификационную таблицу, формулы для расчета чувствительности и специфичности по выборке результатов диагностических тестов выглядят следующим образом:

$$Se \text{ (sensitivity)} = \text{True Positive Rate (TPR)} = A/(A+C) \quad (2)$$

$$Sp \text{ (specificity)} = \text{True Negative Rate (TNR)} = D/(B+D) \quad (3)$$

При наличии матрицы ошибок доступны и могут быть рассчитаны и другие показатели диагностического теста: такие как частота ложноположительных результатов (FPR) или частота промахов, частота ложноотрицательных результатов (FNR), положительное прогностическое значение (PPV), отрицательное прогностическое значение (NPV), коэффициент ложного пропуска (FOR), коэффициент ложного обнаружения (FDR), распространенность или преваленс, доля правильно классифицированных или точных, доля неправильно классифицированных, общая точность, положительное отношение правдоподобия (LR), отрицательное отношение правдоподобия (LR-), отношение диагностических шансов (DOR). Однако в рамках проводимого конкурса они не используются.

Для решения задач конкурса необходим инструмент для оценки и сравнения точности двух или более диагностических алгоритмов, представленных участниками. Это можно сделать путем сравнения чувствительности и/или специфичности с помощью статистических тестов. Ввиду вероятностного характера специфичности и чувствительности оцениваются лишь их статистики, которые в дальнейшем обозначим как:

$$\hat{Se} = \frac{T1}{n1} \quad \hat{Sp} = \frac{F0}{n0} \quad (5)$$

где:

- $T1$ – количество больных (1) диагностированных верно (*True*);
- $n1$ – общее количество больных (1) в исследуемой выборке N ;
- $F0$ – количество ‘не больных’ (0) диагностированных как ‘не больные’ (*False*);
- $n0$ – общее количество ‘не больных’ (0) в исследуемой выборке N .

Если по ним заданы требуемые уровни технологических барьеров, то условием прохождения полуфинала и выходом в финал j -участником конкурса является

преодоление им техбарьера по каждой i -ой нозологии $i=[1;6]$, т.е. выполнение неравенства:

$$Se_i^j > Se^{TP}_i ; Sp_i^j > Sp^{TP}_i \quad (6)$$

Так как Se_i^j , Sp_i^j - вероятностные величины, условием преодоления техбарьера по ним следует считать нижние границы доверительных интервалов для них. Методика их расчета содержится в п.2.

Маловероятно, но теоретически возможно получение одинаковых статистических оценок Se , Sp для диагностики основного заболевания участниками конкурса. В таком случае рассчитывается Se по сопутствующим заболеваниям и участники ранжируются по Se без учета специфичности и расчета доверительных интервалов по ним.

Расчет чувствительности и специфичности ведется с точностью 3-знаков после запятой. Округление 4-го знака осуществляется по правилам математики от 0,0005 в большую сторону.

2. Доверительные интервалы для оценки показателей чувствительности и специфичности

Для статистик Se и Sp формируются доверительные интервалы. В медицинских исследованиях предельно допустимыми по уровню вероятности считаются доверительные интервалы более 0,95. Доверительный интервал (CI) для чувствительности в бинарном диагностическом тесте представляет собой область, в которую попадает истинное значение доли верных предсказаний TP от TP+FN в 95 % случаев. Другими словами, можно с 95 % надежностью сказать, что истинное значение чувствительности диагностического теста в генеральной совокупности будет находиться в пределах 95 % доверительного интервала. Доверительный интервал для специфичности определяет границы диапазона верной диагностики отсутствия, интервал в котором истинное значение специфичности диагностического теста будет находиться с вероятностью не менее 0,95.

Была обоснована возможность и целесообразность расчета доверительных интервалов Se и Sp по непараметрическому методу Уилсона. По мнению многих биомедицинских статистиков, оптимальную оценку доверительных интервалов для частот осуществляет метод Уилсона (Wilson).

Оценка по Уилсону позволяет оценить доверительные интервалы, как для очень малых и очень больших частот, но и применим для малого числа наблюдений.

Для сокращения записей в расчетных формулах для Se и Sp будем использовать обозначение P , рассчитываемую как доля r — количества верных результатов к n - общему числу результатов.

$$\hat{p} = r/n \quad (7)$$

Тогда оценка границ доверительных интервалов для (7) по Уилсону представляется формулами (8):

$$\begin{aligned} \text{Lower Limit} &= \frac{(2n\hat{p} + z_{1-\alpha/2}^2 - 1) - z_{1-\alpha/2} \sqrt{z_{1-\alpha/2}^2 - \{2 + (1/n)\} + 4\hat{p}\{n(1-\hat{p}) + 1\}}}{2(n + z_{1-\alpha/2}^2)} \\ \text{Upper Limit} &= \frac{(2n\hat{p} + z_{1-\alpha/2}^2 + 1) + z_{1-\alpha/2} \sqrt{z_{1-\alpha/2}^2 - \{2 - (1/n)\} + 4\hat{p}\{n(1-\hat{p}) - 1\}}}{2(n + z_{1-\alpha/2}^2)} \end{aligned} \quad (8)$$

Таким образом при расчете доверительные интервалы для чувствительности Se определяются по формуле:

$$\frac{\hat{S}e + \frac{z_{1-\alpha/2}^2}{2nI} \pm z_{1-\alpha/2} \sqrt{\frac{\hat{S}e(1-\hat{S}e) + \frac{z_{1-\alpha/2}^2}{4nI}}{nI}}}{1 + \frac{z_{1-\alpha/2}^2}{nI}} \quad (9)$$

Доверительные интервалы для специфичности Sp:

$$\frac{\hat{S}p + \frac{z_{1-\alpha/2}^2}{2n0} \pm z_{1-\alpha/2} \sqrt{\frac{\hat{S}p(1-\hat{S}p) + \frac{z_{1-\alpha/2}^2}{4n0}}{n0}}}{1 + \frac{z_{1-\alpha/2}^2}{n0}} \quad (10)$$

3. Количество итераций и учет ресурсоемкости в Интегральном критерии

В финальном этапе конкурса для оценки диагностики участнику на вход модели подается файл **xml_v3**, содержащий ограниченный перечень лабораторных и инструментальных исследований по сравнению с версией **xml_v2**.

Так как формат разработанных секций в файлах типов **xml_v3,v2,v1** интероперабелен с информационной моделью ЕГИСЗ для каждого исследования по его коду **NMU** можно получить стоимость, ресурсоемкость по любой диагностической процедуре включенной в датасет (Таблица 1). **Данные таблицы получены согласно Приложению 1 Номенклатуры работ и услуг в здравоохранении (утв. Минздравсоцразвития РФ 12.07.2004)**

Таким образом, последовательно выдавая результаты исследований участнику конкурса можно оценить ресурсоемкость (стоимость, трудоемкость), которую он потратил на диагностику.

Таблица 1. Фрагмент справочника лабораторных исследований

LOIN	FULLNAME	MEASUREMENT	UNIT	SPECIMEN	NMU
35669-1	Амикацин, массовая концентрация в сыворотке или плазме	Массовая концентрация	нг/мл	Сыворотка крови;	Г A09.05.035
78786-1	Амитриптилин, массовая концентрация в крови	Массовая концентрация	мкг/мл	Кровь венозная	A09.05.035
3333-2	Амитриптилин, массовая концентрация в сыворотке или плазме	Массовая концентрация	мкг/мл	Сыворотка крови;	Г A09.05.035
10919-9	Антитела IgE к Anisakis, концентрация в условных единицах	Условные единицы в объёме	кМЕ/л; кЕд/л	Сыворотка крови	A09.05.118
3354-8	Амфотерицин, массовая концентрация в сыворотке или плазме	Массовая концентрация	нг/мл	Сыворотка крови;	Г A09.05.035
16370-9	Амфотерицин, массовая концентрация в моче	Массовая концентрация	нг/мл	Моча	A09.28.016
33915-0	Анабазин, массовая концентрация в моче	Массовая концентрация	мкг/л	Моча	A09.28.016
3364-7	Феназон, массовая концентрация в сыворотке или плазме	Массовая концентрация	мкг/л	Сыворотка крови;	Г A09.05.035
38893-4	Арипипразол, массовая концентрация в сыворотке или плазме	Массовая концентрация	мкг/мл	Сыворотка крови;	Г A09.05.035
78799-4	Атенолол, массовая концентрация в крови	Массовая концентрация	нг/мл	Кровь венозная	A09.05.035
3366-2	Атенолол, массовая концентрация в сыворотке или плазме	Массовая концентрация	нг/мл	Сыворотка крови;	Г A09.05.035
3361-3	Атенолол, массовая концентрация в моче	Массовая концентрация	нг/мл	Моча	A09.28.016
78801-8	Атропин, массовая концентрация в крови	Массовая концентрация	мкг/л	Кровь венозная	A09.05.035
3367-0	Атропин, массовая концентрация в сыворотке или плазме	Массовая концентрация	мкг/л	Сыворотка крови;	Г A09.05.035
10840-7	Атропин, массовая концентрация в моче	Массовая концентрация	мкг/л	Моча	A09.28.016
74391-4	Баклофен, массовая концентрация в крови	Массовая концентрация	нг/мл	Кровь венозная	A09.05.035
9353-4	Баклофен, массовая концентрация в сыворотке или плазме	Массовая концентрация	нг/мл	Сыворотка крови;	Г A09.05.035
75227-9	Баклофен, массовая концентрация в моче	Массовая концентрация	нг/мл	Моча	A09.28.016
3405-8	Бромазепам, массовая концентрация в сыворотке или плазме	Массовая концентрация	мкг/мл	Сыворотка крови;	Г A09.05.035
6706-6	Бупропион, массовая концентрация в сыворотке или плазме	Массовая концентрация	мкг/мл	Сыворотка крови;	Г A09.05.035
9356-7	Бупирон, массовая концентрация в сыворотке или плазме	Массовая концентрация	мкг/мл	Сыворотка крови;	Г A09.05.035
12394-3	Буторфанол, массовая концентрация в сыворотке или плазме	Массовая концентрация	мкг/мл	Сыворотка крови;	Г A09.05.035

Если обозначить:

- D – нозологии конкурса $d = 1, \dots, 6$;
- L_d – перечень лабораторных исследований в датасете для d-ой нозологии
- C_d^L – стоимость лабораторного исследования для d-ой нозологии;
- I_d – перечень инструментальных исследований, присутствующие в датасете для d-ой нозологии;
- C_d^I – стоимость инструментального исследования для d-ой нозологии;

$$C = C_d^L + C_d^I = \sum_n \sum_d C_d^n(I_nmu) + \sum_n \sum_d C_d^n(L_nmu) \quad (11)$$

Накапливая таким образом стоимость диагностики, можно получить оценку стоимости для каждого участника

$$C += C_d^L + C_d^I \quad (12)$$

4. Алгоритм определения победителей

4.1. Полуфинальный этап

Дано: Датасеты **xml_v2**, **xml_v3** по 6-и нозологиям конкурса.

Технологические барьеры Se^{TP_i} и Sr^{TP_i} для каждой нозологии

K – количество Команд участвующих в конкурсе

Каждый уникальный экземпляр в датасет **xml_v2** отличается от аналогичного в **xml_v3**, только тем, что содержит полный набор лабораторных и диагностических исследований пациента. В **xml_v3** – ограниченный набор данных (исследований) для постановки диагноза.

1. Для каждого участника конкурса $k = [1;K]$ выполняем последовательно алгоритм, изображенный на рисунке 2.

Для этого на i -ом шаге на вход СППВР участника передаем **xml_v3** содержащий неполный набор исследований. СППВР выполняет диагностику на неполном наборе и выдает результат.

Если диагностика проведена верно, то считается, что достигнут верный результат $TP_i=1$ или $TN_i=1$ (в зависимости от того больной/небольной пациент на этом шаге) с минимальной ресурсоемкостью $C_i = C_{\min}$.

В случае неверной диагностики по ограниченному **xml_v3**, на вход СППВР подается полный **xml_v2**.

Если диагностика проведена верно, то считается, что достигнут верный результат $TP_i=1$ или $TN_i=1$ (в зависимости от того больной/небольной пациент на этом шаге), но с максимальной ресурсоемкостью $C_i = C_{\max}$.

Если диагностика выполнена неверно, то считается, что верный результат не достигнут $TP_i = 0$ или $TN_i = 0$ (в зависимости от того больной/небольной пациент на этом шаге), и увеличиваются соответствующие значения $FP_i = 1$ или $FN_i = 1$, и ресурсоемкость составила $C_i = C_{\max}$.

Полученные значения на этом шаге TP_i , TN_i , FP_i , FN_i накапливаются в матрице ошибок для соответствующей нозологии.

$$\begin{aligned} TP &+= TP_i \\ TN &+= TN_i \\ FH &+= FN_i \\ FP &+= FP_i \end{aligned} \quad (13)$$

Полученная ресурсоемкость добавляется к накапливаемой ресурсоемкости диагностики СППВР Команды.

$$C += C_i \quad (14)$$

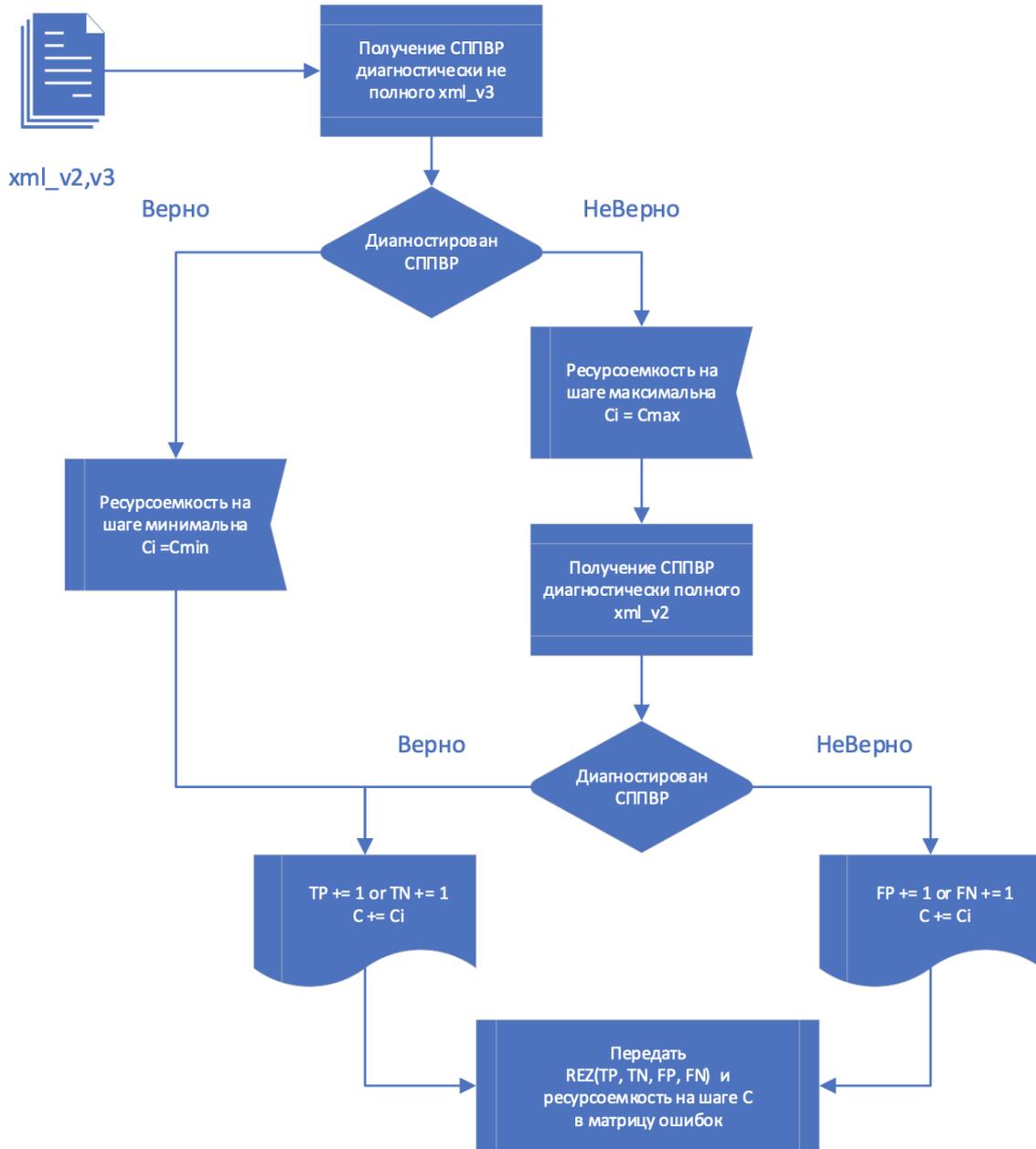


Рис.2 Алгоритм оценки точности и ресурсоемкости диагностики заболевания для нозологии конкурса

2. Для каждой нозологии $i=[1;6]$ и для каждого k -го Участника конкурса $k = [1;K]$ проводим испытание его СППВР. Выдавая результат диагностики, СППВР указывает принадлежность результата к тому или иному классу (болен/неболен).

Строим матрицу ошибок:

Нозология	i	$i=[1,,,6]$
-----------	-----	-------------

Результаты теста по нозологии i		СППВР		
		болен	не болен	всего
Факт	болен	TP 69 (A)	FN 6 (C)	75 ($n1=A+C$)
	не болен	FP 5 (B)	TN 40 (D)	45 ($n0=B+D$)
	всего	72	48	120 (N)

- Для каждой нозологии $i = [1:6]$ рассчитываем показатели чувствительности и специфичности по формулам (2,3);
- Для каждой метрики и по каждой нозологии рассчитываем границы доверительных интервалов CI_{min} ; CI_{max} по формулам (9-10). За уровень значимости принимаем 95% вероятность попадания полученных статистик для каждой метрики в границы интервала $[CI_{min}; CI_{max}]$
- Результаты расчетов каждой метрики i -ой нозологии сохраняем в таблице

Таблица 1. Результаты статистической оценки метрик для i -ой нозологии в границах доверительных интервалов 95%

N п.п.	Метрика	Значение	Доверительный интервал		Технологический барьер	Стоимость ресурсоемкость (Усл.ед)	Результат СППВР
			CI_min	CI_max			
1	Se (TPR)	92,00%	83,73%	96,25%	82,00%	7,968	преодолен
2	Sp (TNR)	88,89%	76,66%	95,12%	79,00%		не преодолен

- Если показатели специфичности и чувствительности равны, рассчитываем **Se** по осложнениям вызванным основным заболеванием и ранжируем по **Se** без учета специфичности и расчета доверительных интервалов по ним.
- Если показатели чувствительности для сопутствующих заболеваний равны, рассчитываем **Se** сопутствующим заболеваниям и ранжируем по **Se** без учета специфичности и расчета доверительных интервалов по ним.

Конечные результаты испытаний по всем нозологиям конкурса переносим в таблицу 2.

Таблица 2. Результаты испытаний СППВР в полуфинале Конкурса

Нозология Конкурса						
	Туберкулез	Пне вмо ния бак	Пне вмо ния вир	Рак легких	ХО БЛ	ХТЭЛГ
Технологические барьеры конкурса по чувствительности и специфичности						
Se Sp	$Se^{TP}_1; Sp^{TP}_1$	$Se^{TP}_6; Sp^{TP}_6$
Результаты испытаний участников						
1	$Se^1_1; Sp^1_1$ $CI_{se}^1 = [CI_{se}^{min 1}; CI_{se}^{max 1}]$; $CI_{sp}^1 = [CI_{sp}^{min 1}; CI_{sp}^{max 1}]$; $Cost^1_1$	$Se^1_N; Sp^1_N$ $CI_{se}^1 = [CI_{se}^{min 1}; CI_{se}^{max 1}]$; $CI_{sp}^1 = [CI_{sp}^{min 1}; CI_{sp}^{max 1}]$; $Cost^1_N$
k	$Se^k_1; Sp^k_1$ $CI_{se}^k = [CI_{se}^{min k}; CI_{se}^{max k}]$; $CI_{sp}^k = [CI_{sp}^{min k}; CI_{sp}^{max k}]$; $Cost^k_1$	$Se^j_k; Sp^j_k$ $CI_{se}^j = [CI_{se}^{min j}; CI_{se}^{max j}]$; $CI_{sp}^j = [CI_{sp}^{min j}; CI_{sp}^{max j}]$; $Cost^j_k$...	$Se^k_N; Sp^k_N$ $CI_{se}^k = [CI_{se}^{min k}; CI_{se}^{max k}]$; $CI_{sp}^k = [CI_{sp}^{min k}; CI_{sp}^{max k}]$; $Cost^k_N$
K	$Se^K_1; Sp^K_1$ $CI_{se}^K = [CI_{se}^{min K}; CI_{se}^{max K}]$; $CI_{sp}^K = [CI_{sp}^{min K}; CI_{sp}^{max K}]$; $Cost^K_1$	$Se^K_N; Sp^K_N$ $CI_{se}^K = [CI_{se}^{min K}; CI_{se}^{max K}]$; $CI_{sp}^K = [CI_{sp}^{min K}; CI_{sp}^{max K}]$; $Cost^K_N$

Если в качестве критериев преодоления технологических барьеров принимаются только статистики по чувствительности и специфичности, то из таблицы 1 на финальную часть отбираются только те СППВР, у которых нижняя граница доверительного интервала преодолевает заданный уровень технологического барьера.

Для всех статистик Se и Sp должны выполняться требования по преодолению нижней расчетной границей доверительного интервала уровней технологических барьеров конкурса для каждой нозологии $d \in D$:

$D = \{ \text{‘Туберкулез’}; \text{‘Бактериальная пневмония’}; \text{‘Вирусная пневмония’}; \text{‘Рак легких’}; \text{‘ХОБЛ’}; \text{‘ХТЭЛГ’} \}$

$$CI_{Se}^{min k} > Se^{mp}_d ; CI_{Sp}^{min k} > Sp^{mp}_d \tag{15}$$

где $k = [1:K]$; $d = [1:D]$

Если технологический барьер преодолели более 4 участников (в финале по условиям конкурса участвуют не более 4 команд) их ранжирование производится следующим образом.

Каждое заболевание имеет свой установленный технологический барьер по чувствительности Se^{mp}_d .

Каждый участник показывает итоговую чувствительность распознавания заболевания которая оценивается по нижней границе доверительного интервала (Таблица 1) $CI_Se^{min}_d^k$.

Для каждого k-участника по каждому d-заболеванию рассчитываем уровень $Y_{k,d}^k$ преодоления им технологического барьера по каждой нозологии конкурса:

$$Y_{k,d}^k = (CI_Se^{min}_d^k - Se^{mp}_d) / Se^{mp}_d \quad (16)$$

Общий уровень преодоления барьеров участником определяется по сумме уровней преодоления по нозологиям:

$$Y_k = \sum_{d=1}^6 Y_{k,d} \quad (17)$$

По результатам расчетов формируется таблица 3. ранжированная по $Y_{k,d}$, в которой $Y_i > Y_j$ для любого i, j из $k = [1; K]$

Таблица 3. Итоговая таблица с интегральным показателем точности диагностики

	ХТЭЛГ	Рак легкого	Туберкулез	ХОБЛ	Пневмония вирусная	Пневмония бактериальная	Интегральный показатель
<i>Участник 1</i>	$Y_{1.1}$	$Y_{2.1}$	$Y_{3.1}$	$Y_{4.1}$	$Y_{5.1}$	$Y_{6.1}$	Y_1
<i>Участник 2</i>	$Y_{1.2}$	$Y_{2.2}$	$Y_{3.2}$	$Y_{4.2}$	$Y_{5.2}$	$Y_{6.2}$	Y_2
<i>Участник 3</i>	$Y_{1.3}$	$Y_{2.3}$	$Y_{3.3}$	$Y_{4.3}$	$Y_{5.3}$	$Y_{6.3}$	Y_3
<i>Участник 4</i>	$Y_{1.4}$	$Y_{2.4}$	$Y_{3.4}$	$Y_{4.4}$	$Y_{5.4}$	$Y_{6.4}$	Y_4
<i>Участник ...</i>							
<i>Участник K</i>	$Y_{1.k}$	$Y_{2.k}$	$Y_{3.k}$	$Y_{4.k}$	$Y_{5.k}$	$Y_{6.k}$	Y_k

Если технологический барьер преодолевает менее 4 команд:

1. Команды, продемонстрировавшие преодоление Технологического барьера на Полуфинальных Испытаниях, попадают в Финальный этап.

2. Оргкомитет использует пороговые значения для отбора в финал конкурса не менее 4 участников, равное $Se = 0,5$ и $Sr = 0,5$. Интегральный показатель и дальнейшее формирование уровней производится также по формулам (16,17).

4.2. Финал

Пусть в финал конкурса вышли $1 \leq K_{\phi} \leq 4$ Команды.

По результатам испытаний Продуктов на Полуфинальном этапе для каждой ССПВР участников сформированы матрицы ошибок $MO_{\text{пф}} = [TP_{\text{пф}}, TN_{\text{пф}}, FP_{\text{пф}}, FN_{\text{пф}}]$ и накоплена ресурсоемкость $Cost_{\text{пф}}$ по каждой нозологии конкурса.

Проводим испытания ССПВР на Финальном этапе Конкурса по алгоритму, представленному на рисунке 3.

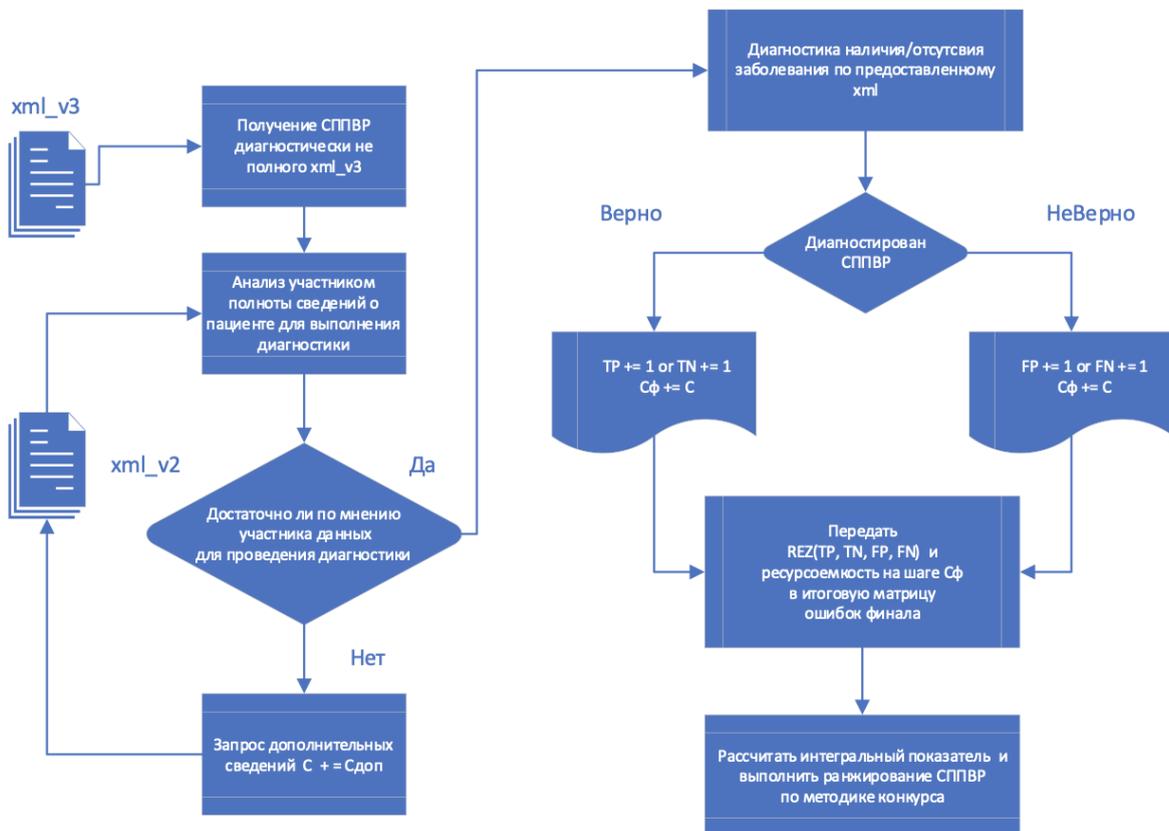


Рис.3 Алгоритм испытаний ССПВР участников в финале

По результатам испытаний полуфинала имеем $K \times 6$ матрицы ошибок MO и для них рассчитаны показатели чувствительности Se , специфичности Sr и нижние границы доверительных интервалов по ним CI_{\min}^{se} , CI_{\min}^{sr} . По итогам финала формируем сводную матрицу ошибок которую к результатам полуфинала добавляем финальные:

$$TP^{\phi} = TP^{\text{пф}} + TP; \quad TN^{\phi} = TN^{\text{пф}} + TN; \quad FP^{\phi} = FP^{\text{пф}} + FP; \quad FN^{\phi} = FN^{\text{пф}} + FN$$

По формулам (5) пересчитываем финальные Se-чувствительность и Sp-специфичность диагностической точности СППВР участников. Рассчитываем для них нижние границы доверительных интервалов (9,10).

N п.п.	Метрика	Значение	Доверительный интервал		Технологический барьер	Стоимость, ресурсоемкость (Усл.ед)	Результат СППВР
			CI_min ^ф	CI_max ^ф			
1	Se ^ф (TPR)	92,00%	84,78%	94,22%	82,00%	257,948	преодолен
2	Sp ^ф (TNR)	87,7%	78,68%	96,12%	79,00%		не преодолен

Если участник не преодолевает Технологический барьер на финальном этапе он исключается из дальнейшего ранжирования для выявления победителя. Для преодолевших формируем итоговую таблицу по каждой нозологии конкурса и определяем интегральный показатель участника для нозологии:

$$H^i = Se_{\phi}^i \times (1 + ((Se^{\max} - Se^{\min}_{\phi}) \times C^{\text{ср}}_{\phi} / C^i_{\phi}))$$

Показатели точности и ресурсоемкости диагностики	Участник 1	Участник 2	Участник 3	Участник 4	Min	Max	Среднее
Чувствительность	Se ¹ _ф	Se ² _ф	Se ³ _ф	Se ⁴ _ф	Se ^{min} _ф	Se ^{max} _ф	Se ^{ср} _ф
Специфичность	Sp ¹ _ф	Sp ² _ф	Sp ³ _ф	Sp ⁴ _ф	Sp ^{min} _ф	Sp ^{max} _ф	Sp ^{ср} _ф
Ресурсоемкость	C ¹ _ф	C ² _ф	C ³ _ф	C ⁴ _ф	C ^{min} _ф	C ^{max} _ф	C ^{ср} _ф
Интегральный показатель	H ¹	H ²	H ³	H ⁴	H ^{max}	H ^{min}	H ^{ср}

На основании полученных значений **H** формируем итоговую таблицу для вычисления Интегрального критерия **I**, и конечного ранжирования участников финала по его значению:

$$I_k = \sum_{d=1}^6 H_{k,d}$$

	ХТЭЛГ	Рак легкого	Туберкулез	ХОБЛ	Пневмония вирусная	Пневмония бактериальная	Интегральный критерий
Участник 1	H _{1.1}	H _{2.1}	H _{3.1}	H _{4.1}	H _{5.1}	H _{6.1}	I ₁
Участник 2	H _{1.2}	H _{2.2}	H _{3.2}	H _{4.2}	H _{5.2}	H _{6.2}	I ₂
Участник 3	H _{1.3}	H _{2.3}	H _{3.3}	H _{4.3}	H _{5.3}	H _{6.3}	I ₃
Участник 4	H _{1.4}	H _{2.4}	H _{3.4}	H _{4.4}	H _{5.4}	H _{6.4}	I ₄

4.2.1 Частичное преодоление технологического барьера

Если ни одна из Команд в процессе финальных Испытаний не преодолевает Технологический барьер - используется методика определения частичного преодоления. СППВР ранжируются по значению Интегрального критерия **И**. Частично преодолевшими технологический барьер считаются от одной до трех Команд, чьи СППВР набрали наивысшее значение Интегрального критерия и при этом преодолели пороговые значения $Se = 0,7$ и $Sr = 0,7$ по каждой из нозологий Конкурса по результатам Финального этапа.

Приложение 2. Типы Сессий обмена с Платформой в рамках Отборочного этапа Конкурса

Тренировочная Сессия

Тип Сессии - training

Ответ должен приходиться в виде массива объектов, каждый объект которого должен включать в себя следующие ключи:

- XPath - со значением "" (пустая строка)
- start - со значением "" (пустая строка)
- end - со значением "" (пустая строка)
- name - текстовое представление шифра МКБ-10 из Эпикриза
- decorCode - со значением "diagnosisMain"
- code - со значением "" (пустая строка)

Алгоритмическая Сессия (для прохождения Алгоритмической квалификации)

Тип Сессии - algorithmic

Ответ должен приходиться в виде массива объектов, каждый объект которого должен включать в себя следующие ключи:

- XPath - со значением "" (пустая строка)
- start - со значением "" (пустая строка)
- end - со значением "" (пустая строка)
- name - текстовое представление шифра МКБ-10 из Эпикриза
- decorCode - со значением "diagnosisMain"
- code - со значением "" (пустая строка)

Ограничение

При проведении Алгоритмической квалификации, участник за одну Сессию будет обрабатывать один Эпикриз. Следовательно ответ должен будет содержать массив с одним объектом. В случае, если участник пришлет несколько объектов в массиве - система автоматически примет такой ответ как неправильный, вследствие чего Алгоритмическая квалификация будет считаться не пройденной.