

ПРОЕКТ

Технологический конкурс в целях реализации Национальной
технологической инициативы

AI'M Doctor

по созданию интеллектуальной системы поддержки принятия врачебных
решений

Технический регламент

1.0

Данная версия технического регламента предназначена для получения обратной связи от потенциальных участников технологического Конкурса AI'M Doctor и не является окончательной

Положения данной версии регламента могут быть изменены организаторами в любой момент времени

Оглавление

1. Общие положения	3
2. Термины и определения	3
3. Условия участия в конкурсе	6
4. Структура Конкурса	6
5. Конкурсная задача и методика оценивания Решений участников	7
6. Общие требования к процессам Испытаний на всех этапах	9
7. Отборочный этап	10
8. Полуфинальный этап	12
9. Финальный этап	14
10. Судейство Конкурса	16
11. Ограничения и дисквалификации	16
12. Обеспечение безопасности Конкурса	17
13. Дополнительные условия	18
Приложение 1. Интегральный критерий и метрики оценивания Решений	20

1. Общие положения

- 1.1. Настоящий Технический регламент определяет требования к процедуре проведения технологического конкурса AI'M Doctor в целях реализации Национальной технологической инициативы (далее Конкурс), обеспечению и другим техническим параметрам Конкурса.
- 1.2. Настоящий Технический регламент является документом, детализирующим Конкурсное задание Конкурса, опубликованное на 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.
- 1.3. 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. Термины и определения

- 2.1. **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.
- 2.2. **Быстро возрастающая последовательность** – последовательность вещественных положительных чисел, таких что каждый следующий член последовательности больше суммы всех предыдущих членов последовательности.
- 2.3. **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 совокупностей медицинских данных (эпикризов). Валидационный датасет Финального этапа - **совокупность данных, содержащая валидационный датасет Полуфинального этапа и валидационный датасет финальных испытаний.**

- 2.4. **Демонстрация решения** - мероприятие, в процессе которого производится демонстрация работы представленного решения участников в целях измерения и определения результатов выполнения ими Конкурсной задачи с целью определения Судейской коллегией соответствия демонстрируемого решения поставленным критериям для преодоления технологического барьера.
- 2.5. **Заключительный клинический диагноз** - результат работы СППВР и/или клинического специалиста, врача, который включает в себя: основное заболевание (код МКБ 10), осложнения основного заболевания (список кодов по МКБ 10), сопутствующие заболевания (список кодов по МКБ 10).
- 2.6. **Зачетные испытания, Испытания, Финальные испытания** - период времени, начинающийся в назначенный Оргкомитетом момент, когда Участники проводят с помощью разработанных ими Программных комплексов процесс постановки диагнозов по множеству эпикризов валидационного датасета
- 2.7. **Конкурс** – множество технологических мероприятий, проводимых в целях реализации Национальной технологической инициативы по созданию интеллектуальной системы поддержки принятия врачебных решений. Предметом конкурса является выбор наилучшего решения для автоматической постановки заключительного клинического диагноза.
- 2.8. **Клинический специалист** - лицо, осуществляющее деятельность направленную на комплексную проверку СППВР Участников в рамках финального этапа Конкурса.
- 2.9. **Ознакомительный датасет** - набор эпикризов, предназначенный для демонстрации структуры и вида данных валидационных датасетов (эпикризов) полуфинального и финального этапов. Ознакомительный датасет может содержать все необходимые данные для постановки диагноза и сам диагноз (включая осложнения и сопутствующие заболевания)
- 2.10. **Осложнения основного заболевания** – присоединившийся к основному заболеванию патологический процесс, отягощающий типичное течение болезни; обусловлен не её причиной, а дополнительными изменениями, возникшими в организме в ходе заболевания (напр., воспаление лёгких в послеоперационном периоде).
- 2.11. **Основное заболевание** – заболевание, которое явилось причиной обращения к врачу, лечебно-диагностических мероприятий, а при неблагоприятном исходе может привести к инвалидизации или смерти больного.

- 2.12. **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 Решений в соответствии с математическим аппаратом Интегрального критерия
- 2.13. **Программный комплекс участника (ПКУ)** – разработанный Командой Участника Конкурса комплекс программного и аппаратного обеспечения, который в автоматическом режиме принимает на входе эпикризы и выдаёт на выходе элементы заключительного клинического диагноза по установленной форме
- 2.14. **Решение Команды, Решение** - 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
- 2.15. **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
- 2.16. **Тестовые испытания** - процесс, симулирующий отборочные, полуфинальные или финальные мероприятия и направленный на выявление различных ошибок процесса испытаний или систем СППВР
- 2.17. **Судейская коллегия** - совещательный орган правомочный вести наблюдение за процессами реализации Конкурса на всех этапах, задачей которого является подтверждение соответствия процессов Испытаний Конкурсной документации (Конкурсное задание, Технический регламент), а также фиксация и подтверждение результатов команд. В случае возникновения спорных ситуаций Судейская коллегия принимает итоговое решение.
- 2.18. **Техническая комиссия** - совещательный орган, верифицирующий процессы Испытаний с точки зрения технической корректности. В случае возникновения спорных ситуаций технического характера, препятствующих или вносящих изменения в штатный ход испытаний, Техническая комиссия принимает решение по дальнейшим действиям.
- 2.19. **Эпикриз** - совокупность анонимизированных данных об одном пациенте, содержащая как результаты врачебных осмотров, так и результаты клинических исследований.

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 данным требованиям.
- 3.4. Все Участники, на всех этапах Конкурса, а также при разработке своих ПКУ и получаемых с помощью ПКУ решений 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 участников 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 участников выполнять задачу Конкурса в реальных условиях применения современных интеллектуальных медицинских систем. Полуфинальный этап подробно описан в п.9 настоящего Технического регламента, а также в п.5.4 Конкурсного задания Конкурса.
- 1.1. Каждый этап Конкурса заканчивается Испытаниями. Испытания могут иметь разный формат (быть очными или заочными; проводиться дистанционно или оффлайн). Под дистанционными Испытаниями понимается процесс, когда все команды подключаются одновременно и

работают в параллельном режиме над решением определенной задачей. Ход дистанционных Испытаний контролируется Судейской коллегией и Технической комиссией.

- 4.2. После успешной регистрации участник получает доступ к Платформе. В рамках Платформы участникам доступно:
 - 4.2.1. Личный кабинет с необходимыми исходными данными и результатами прохождения этапов.
 - 4.2.2. Датасеты конкурса (перечень и наполнение датасетов зависит от текущего этапа Конкурса)
 - 4.2.3. 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>
- 4.3. Использование общих программных модулей и установленных разработчиком Платформы способов взаимодействия с Платформой, является обязательным, поскольку обеспечивает равные условия для всех Участников Конкурса.
- 4.4. Участники Конкурса не имеют права самостоятельно вносить изменения в общие программные модули. В случаях обнаружения ошибок в них Участник должен подать в Оргкомитет по электронной почте в свободной форме заявку «О внесении изменений в общие программные модули», описав в ней обнаруженные ошибки и/или свои предложения по улучшению работы модулей.

5. Конкурсная задача и методика оценивания Решений участников

- 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. Технологический барьер - **точность постановки диагноза по основному заболеванию на основе статистических данных**
- 5.4. Для качественного ранжирования результатов Команд в рамках Конкурса используется математический аппарат, обеспечивающий комплексное многофакторное оценивание Решений - Интегральный критерий. Подробное описание метрик и методик оценивания приведено в Приложении 1 настоящего Технического регламента.

- 5.5. Для согласования работы всех субъектов и сущностей Конкурса используется распределенная информационная система Конкурса - Платформа. Все процессы испытаний на всех этапах конкурса подразумевают взаимодействие программных продуктов участников с Платформой Конкурса.
- 5.6. Процессы испытаний подразумевают общую итерационную последовательность действий:
- Шаг 1. Платформа предоставляет набор данных для чтения системами участников
 - Шаг 2. Система участника анализируют полученные данные и формируют результат анализа (в соответствии с требованиями этапа)
 - Шаг 3. Результат анализа передается системой участника на Платформу
 - Шаг 4. Возврат на шаг 1.
- 5.7. Элементы валидационного датасета (эпикризы в виде **XML-файлов** с входными анонимизированными данными пациентов, содержащими жалобы пациента, анамнестические данные, протоколы исследований и консультаций с врачами) поступают на вход СППВР Участников.
- 5.8. Отборочный этап:
- 1.1.1. Платформа передает СППВР 1 случайный эпикриз из Валидационного дата-сета
 - 1.1.2. **СППВР обрабатывает файл и возвращает на платформу код МКБ основного диагноза. В рамках отборочного этапа используются эпикризы, содержащие информацию о диагнозе.**
- 5.9. Полуфинальный этап:
- 5.9.1. В рамках данного этапа задачей СППВР участника является:
 - 5.9.1.1. Постановка предварительного клинического диагноза по каждому из эпикризов Валидационной выборки (на основе неполных данных).
 - 5.9.1.2. Постановка Заключительного клинического диагноза по каждому из эпикризов Валидационной выборки.
 - 5.9.2. 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)
 - 5.9.3. Система проверки решения сравнивает результат работы СППВР конкурсантов с эталонным результатом, предварительно сформированным **экспертной группой врачей.**

- 5.9.4. По результатам сравнения рассчитываются метрики:
 - 5.9.4.1. **точность (чувствительность и специфичность) выявления каждой из нозологий основного диагноза;**
 - 5.9.4.2. точность выявления осложнений и сопутствующих заболеваний.
 - 5.9.4.3. **стоимость постановки диагноза (будут учтены на финальном этапе Конкурса).**

5.10. Финальный этап:

- 5.10.1. В рамках данного этапа системам Участников необходимо:
 - 5.10.1.1. Определить заключительный клинический диагноз по всем эпикризам валидационного датасета.
 - 5.10.1.2. 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.
- 5.10.2. Критерии оценки СППВР включают в себя: точность постановки диагноза; оптимальность постановки диагноза.

6. Общие требования к процессам Испытаний на всех этапах

- 6.1. Для согласования работы всех субъектов и сущностей Конкурса используется распределенная информационная система Конкурса - Платформа. Участие Команды во всех процессах Конкурса возможно и разрешено только с использованием документированного функционала и инструментария Платформы.
- 6.2. Для участия в Полуфинальных и Финальных испытаниях Участник должен запустить сессию обмена данными с Платформой до момента старта испытаний. В случае подключения к Платформе после начала Испытаний, метрика для пропущенных Командой файлов эпикризов аннулируется (**диагноз считается поставленным неверно**).
- 6.3. В рамках Полуфинального и Финального этапов проводится 2 вида Испытаний: Тестовые испытания и Зачетные испытания. Тестовые испытания не являются обязательными для команд и проводятся в той же логике, что и Зачетные испытания, но на меньшей выборке данных. Результаты СППВР на Тестовых испытаниях не оцениваются.
- 6.4. Судейская коллегия, Техническая комиссия и экспертная группа проводит предварительную ревизию Решений на предмет соответствия Решений целям и задачам Конкурса. Результаты предварительной ревизии

передаются в Оргкомитет. В случае, если Решение какой-либо из Команд не соответствует целям и задачам Конкурса - данное Решение аннулируется Оргкомитетом и расчет метрик интегрального критерия не производится.

- 6.5. В течение суток с момента окончания сессии Испытаний в личном кабинете участника формируется отчет, отражающий результаты Команды.
- 6.6. В процессе подготовительных этапов и непосредственно Испытаний Оргкомитетом предоставляется канал коммуникации между командами и организаторами. Информация об организации данного канала коммуникации доводится до участников после регистрации Команды
 - 6.6.1. В качестве канала коммуникации могут использоваться инструменты мессенджеров или сред совместной работы.
 - 6.6.2. Информация публикуемая с использованием канала коммуникации считается официально доведенной до участников.
 - 6.6.3. Участники команд обязаны иметь в канале коммуникации имя пользователя (ник), соответствующий следующей структуре:
Название команды_ФИО участника полностью.
- 6.7. 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:
 - 6.7.1. В качестве даты новых Испытаний может быть назначен другой день (но не далее чем 7 дней с момента отмененного этапа), о чём Оргкомитет уведомляет отдельно.
 - 6.7.2. Используется новый датасет. Все предыдущие результаты команд в рамках данного этапа могут быть аннулированы. Регламент действий в случае перезапуска определяется Оргкомитетом.
- 6.8. Прямое взаимодействие участников команды и членов Судейской коллегии или Технической комиссии в рамках процессов Конкурса допускается только с участием представителя Оргкомитета.

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 режиме.
- 7.4. Отборочный этап состоит из:
 - 7.4.1. Процесса формирования, подачи и одобрения заявки на участие в Конкурсе.
 - 7.4.2. Процесса интеграции Команды с Платформой (получения доступа к личному кабинету пользователя, установки необходимого ПО и т.д.).
 - 7.4.3. Выполнения задания Отборочного этапа
- 7.5. После одобрение Оргкомитетом заявки Команды на участие в Конкурсе, Участник получает доступ в Личный кабинет Участника на Платформе.
- 7.6. Участник должен загрузить с Платформы необходимые программные модули и развернуть клиентскую часть ПКУ для подключения к Платформе собственных систем СППВР.
- 7.7. В рамках Отборочного этапа Участникам становится доступен **Ознакомительный датасет**
- 7.8. Для успешного прохождения Отборочного этапа Команда должна пройти алгоритмическую квалификацию - загрузить задание отборочного этапа, проанализировать его и вернуть на Платформу решение: **основной диагноз в виде кода МКБ**
- 7.9. Задание отборочного этапа может являться как образцом реального анонимизированного эпикриза, так и набором синтетических данных (имеющих структуру реального эпикриза)
- 7.10. Задание Отборочного этапа в своей структуре уже содержит сведения о диагнозе. Таким образом, задача участника заключается в определении этих данных внутри структуры эпикриза, считывании данных и возврате данных на Платформу.
- 7.11. Для проведения алгоритмической квалификации Участник может открыть сессию обмена данными с Платформой в любой момент (**в рамках времени испытаний Отборочного этапа**). При открытии сессии сервер Платформы отправляет ПКУ задание и ожидает ответ от ПКУ.
- 7.12. 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).

- 7.13. Команда, прошедшая алгоритмическую квалификацию, допускается к Полуфинальному этапу.

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 Участнику передается Ознакомительный датасет. Участник вправе использовать Ознакомительный датасет как обучающую выборку, однако, организатор Конкурса не гарантирует репрезентативность и достаточность выборки.
- 8.4. Испытания полуфинального этапа проводятся на основе Валидационного датасета, имеющего объем не менее 500 элементов. Каждый элемент датасета представляет из себя набор **анонимизированных** медицинских данных об одном пациенте (клинические исследования, результаты анализов, жалобы и симптомы, Объективные данные) - эпикриз. Диагнозы в Валидационном датасете Полуфинального этапа отсутствуют.
- 8.5. Каждый элемент датасета представлен в двух версиях - неполные данные и полные данные
- 8.6. 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 всем Участникам Конкурса и исключения любых возможностей для организации ручной постановки диагноза. Расписание Финальных испытаний (дата-время старта и периодичность отдачи файлов) объявляется не менее, чем за 5 (пять) рабочих дней до их начала.
- 8.7. **Полуфинальные испытания проходят в два подэтапа (предварительная постановка диагноза и окончательная постановка диагноза). Процесс реализуется следующим образом:**
- 8.7.1. **Платформа публикует случайный эпикриз (версии неполных данных)**
- 8.7.2. **СППВР Участников анализирует файл, и на основе информации в нем возвращает предполагаемый основной диагноз (в виде кода МКБ). На данную операцию отводится 30 секунд.**
- 8.7.3. **Платформа публикует полную версию эпикриза**
- 8.7.4. **СППВР Участников анализирует файл, и на основе информации в нем возвращает:**
 - либо код МКБ Основного заболевания (одна из шести утвержденных нозологий). Коды МКБ Сопутствующих

заболеваний и Осложнений;
 - либо значение “Не болен”, если пациент не болен ни одной из нозологий Конкурса. На эту операцию отводится 30 секунд.

- 8.8. 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 данного Участника.
- 8.9. Если аннулирование файла эпикризов происходит у нескольких Участников по причине 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 Участников.
- 8.10. По окончании циклов отправки/приема всего объема выборки Полуфинального датасета сессии обмена данными между ПКУ и Платформой закрываются, все данные, полученные от Участников, сохраняются в Базу данных диагнозов (формируя закрытый датасет с полуфинальными решениями Участников). С этого момента стартует процесс оценивания решений Участников.
- 8.11. После предварительной ревизии решений запускается процесс расчета Интегрального критерия по выборке решений Команд
- 8.12. По итогам Полуфинального этапа строится лидерборд — список Участников, ранжированный по убыванию интегрального критерия (см Приложение 1). Аннулированные решения не участвуют в финальном лидерборде.
- 8.13. 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.
- 8.14. Преодоление технологического барьера на Полуфинальном этапе Конкурса не является условием победы команды в Конкурсе и определяет только возможность к переходу на Финальный этап.
- 8.15. Если по результатам Полуфинального этапа не может быть сформирован пул из 4 участников 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. Финальный этап начинается с момента подведения итогов Полуфинального этапа и завершается очными Испытаниями с личным присутствием Участника на площадке Испытаний.
- 9.3. Финальные испытания проводятся на основе выборки Финального датасета, имеющего объем не менее 16 элементов. Каждый элемент датасета представляет из себя набор **анонимизированных** медицинских данных об одном пациенте (клинические исследования, результаты анализов, жалобы и симптомы, Объективные данные) - эпикриз. Диагнозы в Валидационном датасете финального этапа отсутствуют.
- 9.4. 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.
- 9.5. Финальный этап представляет собой процесс последовательной обработки элементов Финального датасета (эпикризов пациентов) с целью постановки основного клинического диагноза, определения осложнений и сопутствующих заболеваний.
- 9.6. Обработка каждого элемента Финального датасета выполняется в 2 шага:
 - 9.6.1. Шаг А: Итерационная постановка диагнозов в режиме ограниченного времени.
 - 9.6.1.1. СППВР получает первичные **анонимизированных** данные о пациенте и анализирует данные. Набор и вид первичных **анонимизированных** данных определяется после завершения Полуфинального этапа Конкурса
 - 9.6.1.2. 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.
 - 9.6.1.3. СППВР запрашивает дополнительную информацию о пациенте. За одну итерацию можно запросить результаты только одного анализа или исследования.
 - 9.6.1.4. 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.9.6.1.2

9.6.1.5. Шаги 9.5.1.1 - 9.5.1.3 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).

9.6.1.6. На Платформе сохраняется информация о траектории постановки диагнозов и комментарии Клиническому специалисту.

9.6.2. Шаг Б: Работа в паре с Клиническим специалистом

9.6.2.1. Определение пар СППВР-Клинический специалист происходит по методике п.5.4.17 Конкурсного задания Конкурса.

9.6.2.2. Взаимодействие СППВР и Клинического специалиста осуществляется через Платформу.

9.6.2.3. Платформа воспроизводит работу СППВР по траектории, сохраненной на шаге А.

9.6.2.4. 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".

9.7. После окончания Испытаний решения Команд и Клинических специалистов сохраняются на Платформе, формируя закрытый датасет результатов Финального этапа.

9.8. Платформа производит расчет метрик интегрального критерия на выборке данных закрытого датасета результатов Финального этапа.

9.9. Испытания финального этапа проводятся очно на территории Организатора. В процессе проведения Финальных испытаний на площадке испытаний должны присутствовать минимум двое представителей команды.

9.10. Процесс Финальных испытаний регламентируется организационными правилами, которые формируются Оргкомитетом Конкурса и доводятся до сведения Участников не менее 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 Участников данных, выявляет условия и причины аннулирований на стороне Платформы, если таковые имеются.
- 10.4. Прямое взаимодействие Участников с членами Судейской коллегии и Технической комиссии может происходить только с прямого разрешения Оргкомитета и в присутствии представителей Оргкомитета.
- 10.5. По окончании Финальных испытаний и процедур расчета интегрального критерия Судейская коллегия проверяет автоматически формируемые отчеты о решениях Участников и условия преодоления технологического барьера.
- 10.6. Процесс определения победителей подробно описан в п.6 и п.7 Конкурсного задания Конкурса.
- 10.7. Оргкомитет выносит итоговые решения по результатам Полуфинального и Финального этапов, которые передаются на рассмотрение и утверждение Жюри.

11. Ограничения и дисквалификации

- 11.1. Участники Конкурса обязаны следовать требованиям Конкурсного Задания и Технического регламента Конкурса
- 11.2. В процессе Конкурса Участники обязаны следовать принципам честности и нормам морали принятым в обществе при взаимодействии с организаторами или другими участниками, соблюдать требования этики, быть внимательным и доброжелательным в общении с другими участниками и организаторами, поддерживать имидж организатора, воздерживаться от некорректных высказываний в отношении профессионализма других участников, воздерживаться от оскорблений в адрес других участников, организаторов и партнеров Конкурса.
- 11.3. Участнику 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 механикой Технического регламента и Конкурсного задания (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, участнику запрещается использовать данные особенности в целях дискредитации Конкурса или получения выгоды.
- 11.9. Запрещается использовать логические, алгоритмические и иные решения, противоречащие цели и задачам Конкурса. В случае подозрения в использовании подобных решений Оргкомитет имеет право дисквалифицировать Команду. Настоящее ограничение распространяется и на особенности системы, 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 Участника 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 участниками) 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 Участникам правила по технике безопасности и охране окружающей среды, с которыми все члены команды Участника 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 Участникам приходит сообщение «Пауза», после чего сервер Платформы временно перестает выполнять штатные действия по реализации испытаний. По окончании паузы всем Участникам 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.
- 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 влияют на сутевую составляющую Конкурсного задания и Технического регламента, 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 Участников 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 Решений

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');

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

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

Таким образом если исследуется *N* пациентов, то *A* и *D* – количества верных диагностических ответов, а *B* и *C* – количества неверных диагнозов, 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). Иными словами, чувствительность диагностического теста — это доля тех, у кого есть состояние, для которого диагностический тест положительный, а специфичность диагностического теста — это доля тех, у кого нет состояния, для которого диагностический тест является положительным, отрицательный. 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.

2. Доверительные интервалы для оценки показателей чувствительности и специфичности

Для статистик Se и Sp необходимо сформировать 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.

Ввиду того, что характер распределения этих случайных величин нам неизвестен мы не можем применять традиционный подход, который использует нормальное приближение к биномиальному распределению, а воспользуемся методами непараметрической оценки.

Ранее на стратегической сессии конкурса 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. Нами была проведена сравнительная оценка интервалов по Уилсону с иными методами для диагностической бинарной классификации в результате чего установлено, что оценка по Уилсону демонстрирует более высокую точность чем биномиальное распределение или доверительные интервалы по Вальду.

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:

$$\frac{\hat{S}e + \frac{z_{1-\alpha/2}^2}{2n1} \pm z_{1-\alpha/2} \sqrt{\frac{\hat{S}e(1-\hat{S}e) + \frac{z_{1-\alpha/2}^2}{4n1}}{n1}}}{1 + \frac{z_{1-\alpha/2}^2}{n1}} \quad (9)$$

Доверительные интервалы для специфичности Sp:

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

При необходимости, для любой метрики оценки из (4) (кроме LR) по формулам (7)-(10) могут быть рассчитаны доверительные интервалы.

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_v3**.

Так как формат разработанных секций в файлах **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).

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 объеме	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,..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;

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

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

$$C += C_d^L + C_d^I \tag{12}$$

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 Sp^{TP_i} для каждой нозологии

K – количество команд участвующих в конкурсе

Каждый уникальный экземпляр в датасет **xml_v3** отличается от аналогичного в **xml_v2**, только тем, что содержит полный набор лабораторных и диагностических исследований пациента. В **xml_v3** – указываются не все, но в случае правильной постановки диагноза по ограниченному набору данных.

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.

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

Полученная ресурсоемкость добавляется к накапливаемой ресурсоемкости диагностики СППВР участника.

$$C += C_i \tag{14}$$

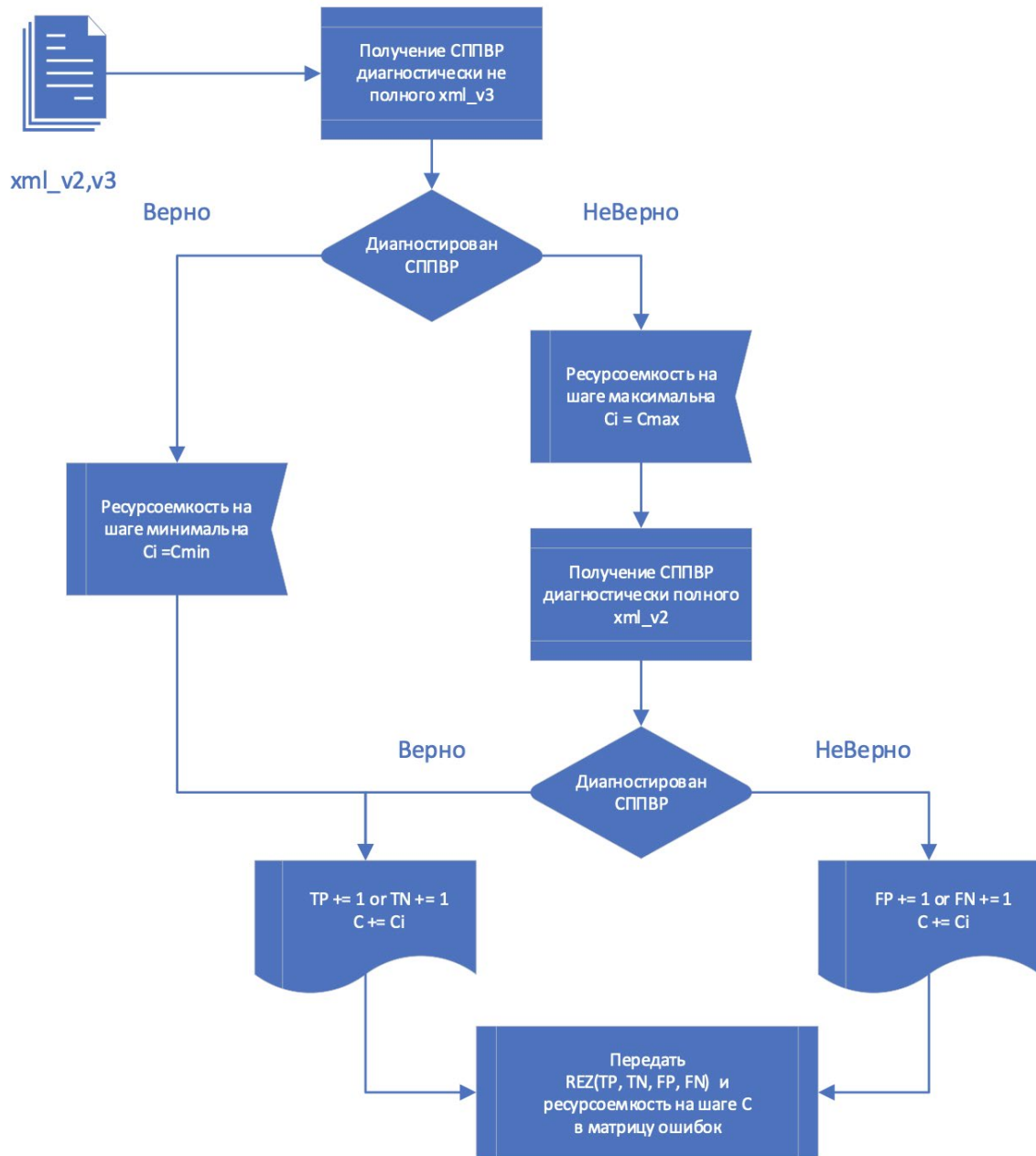


Рис.2 Алгоритм оценки точности и ресурсоемкости диагностики заболевания для нозологии конкурса

- Для каждой нозологии $i=[1;6]$ и для каждого k -го участника конкурса $k = [1;K]$ проводим испытание его СППВР. Выдавая результат диагностики, участник 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 вызванными 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	$Se^1_1; Sp^1_1$ $CI_{Se} = [CI_{Se}^{min_1 1}; CI_{Se}^{max_1 1}]$; $CI_{Sp} = [CI_{Sp}^{min_1 j}; CI_{Sp}^{max_1 j}]$; $Cost^1_1$	$Se^1_N; Sp^1_N$ $CI_{Se}^1_N = [CI_{Se}^{min_N 1}; CI_{Se}^{max_N 1}]$; $CI_{Sp}^1_N = [CI_{Sp}^{min_N 1}; CI_{Sp}^{max_N 1}]$; $Cost^1_N$
...
k	$Se^k_1; Sp^k_1$ $CI_{Se}^k_1 = [CI_{Se}^{min_1 k}; CI_{Se}^{max_1 k}]$; $CI_{Sp}^k_1 = [CI_{Sp}^{min_1 k}; CI_{Sp}^{max_1 k}]$; $Cost^k_1$	$Se^j_k; Sp^j_k$ $CI_{Se} = [CI_{Se}^{min_k j}; CI_{Se}^{max_k j}]$; $CI_{Sp} = [CI_{Sp}^{min_k j}; CI_{Sp}^{max_k j}]$; $Cost^k_j$...	$Se^k_N; Sp^k_N$ $CI_{Se}^k_N = [CI_{Se}^{min_N k}; CI_{Se}^{max_N k}]$; $CI_{Sp}^k_N = [CI_{Sp}^{min_N k}; CI_{Sp}^{max_N k}]$; $Cost^k_N$
...
K	$Se^K_1; Sp^K_1$ $CI_{Se}^K_1 = [CI_{Se}^{min_1 K}; CI_{Se}^{max_1 K}]$; $CI_{Sp}^K_1 = [CI_{Sp}^{min_1 K}; CI_{Sp}^{max_1 K}]$; $Cost^K_1$	$Se^K_N; Sp^K_N$ $CI_{Se}^K_N = [CI_{Se}^{min_N K}; CI_{Se}^{max_N K}]$; $CI_{Sp}^K_N = [CI_{Sp}^{min_N K}; CI_{Sp}^{max_N K}]$; $Cost^K_N$

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 $d \in D$:

$D = [\text{'Туберкулез'}; \text{'Бактериальная пневмония'}; \text{'Вирусная пневмония'}; \text{'Рак легких'}; \text{'ХОБЛ'}; \text{'ХТЭЛГ'}]$

$$CI_{Se}^{min_d k} > Se^{mp_d} \quad ; \quad CI_{Sp}^{min_d k} > Sp^{mp_d} \quad (15)$$

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 преодоления им технологического барьера по каждой нозологии конкурса:

$$Y^k_d = (CI_Se^{min_d^k} - Se^{mp_d}) / Se^{mp_d} \quad (16)$$

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_{kd} , 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
Участник 1	$Y_{1,1}$	$Y_{2,1}$	$Y_{3,1}$	$Y_{4,1}$	$Y_{5,1}$	$Y_{6,1}$	Y_1
Участник 2	$Y_{1,2}$	$Y_{2,2}$	$Y_{3,2}$	$Y_{4,2}$	$Y_{5,2}$	$Y_{6,2}$	Y_2
Участник 3	$Y_{1,3}$	$Y_{2,3}$	$Y_{3,3}$	$Y_{4,3}$	$Y_{5,3}$	$Y_{6,3}$	Y_3
Участник 4	$Y_{1,4}$	$Y_{2,4}$	$Y_{3,4}$	$Y_{4,4}$	$Y_{5,4}$	$Y_{6,4}$	Y_4
Участник ...							
Участник K	$Y_{1,k}$	$Y_{2,k}$	$Y_{3,k}$	$Y_{4,k}$	$Y_{5,k}$	$Y_{6,k}$	Y_k

Если технологический барьер преодолет менее 4 команд, то организаторы конкурса снижают технологические барьеры Se^{mp_d} до уровня достаточного для отбора в финал конкурса не менее 4 участников, но не ниже $Se = 0,7$ и $Sr = 0,65$. Интегральный показатель и дальнейшее формирование уровней производятся также по формулам (16,17).

4.2. Финал

Пусть в финал конкурса вышли $1 \leq K_{\phi} \leq 4$ участника и они преодолели технологические барьеры по каждой нозологии конкурса.

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

Проводим испытания СППВР участников на финальном этапе конкурса по алгоритмы представленному на рисунке 3.

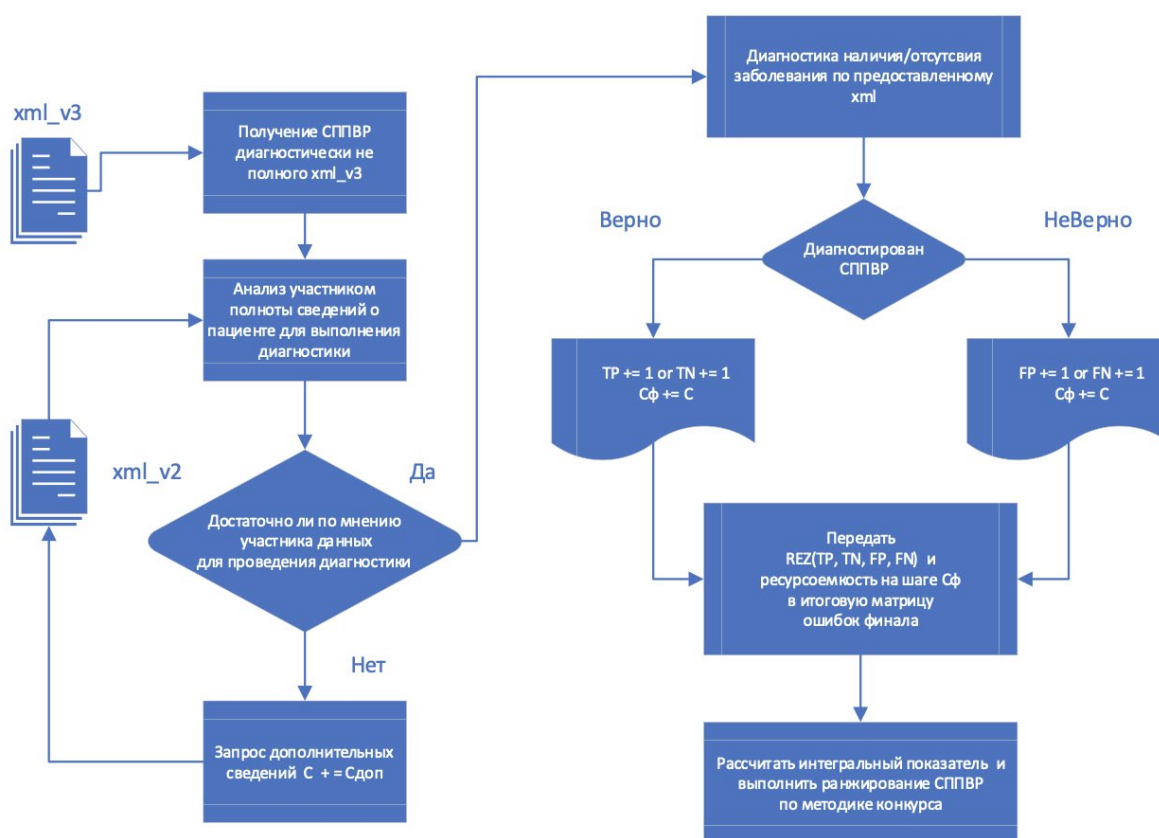


Рис.3 Алгоритм испытаний СППВР участников в финале

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

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

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 его исключаем из дальнейшего ранжирования для выявления победителя. Для преодолевших формируем итоговую таблицу по каждой нозологии конкурса и определяем интегральный показатель участника для нозологии:

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

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

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

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

Победителем признается k-участник с наибольшим значением показателя – **I_k**

$$\max \{I_k\} \quad k = 1, \dots, 4$$

	ХТЭЛГ	Рак легкого	Туберкулез	ХОБЛ	Пневмония вирусная	Пневмония бактериальная	Интегральный показатель
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

4.3. Технологические барьеры нозологий:

Нозология	Рак легкого	Туберкулез	Пневмония	ХТЭЛГ	ХОБЛ
Чувствительность	78,69%	79,625%	69%	81,56%	75%
Специфичность	89,72%	88,5%	79,5%	93,84%	84%