

Технологический конкурс в целях реализации Национальной технологической инициативы

AI'm Doctor

по созданию интеллектуальной системы поддержки принятия врачебных
решений

Технический регламент

1.4

Оглавление

1. Термины и определения	3
2. Общие положения	5
3. Условия участия в конкурсе	6
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	12
8. Полуфинальный этап	13
9. Финальный этап	16
10. Конкурсная площадка	20
11. Судейство Конкурса	21
12. Ограничения и дисквалификации	22
13. Особые случаи и обеспечение безопасности Конкурса	23
14. Дополнительные условия	23
Приложение 1а	25
Метрики оценивания Совокупностей данных	25
Приложение 1б.	31
Алгоритм определения победителей Полуфинального этапа	31
Приложение 1в	37
Алгоритм определения победителей Финального этапа	37
Методика определения Частичного преодоления технологического барьера и ранжирования Команд:	40
Решение: Ранжирование по точности без учета нозологий. Два призовых места.	42
Приложение 2	43
Типы Сессий обмена с Платформой в рамках Отборочного этапа Конкурса	43
Приложение 3	44
Типы Сессий обмена с Платформой в рамках Полуфинального этапа Конкурса	44
Общая информация	44
Формат ответа	44
Типы сессий:	45
Тренировочная сессия	45
Расчетная тренировочная сессия	46
Испытательная сессия	47
Примеры ответов	47
Приложение 4	
Таблица исследований Финального этапа	50

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. **Интерпретируемость** - способность СППВР корректно (с точки зрения Клинического специалиста) формировать предположение о диагнозе по набору первичных данных, формировать на основе данного предположения траекторию постановки (подтверждения) диагноза, ставить заключительный клинический диагноз.
- 1.6. **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.
- 1.7. **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.
- 1.8. **МКБ 10** - Международная классификация болезней 10-го пересмотра

- 1.9. **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 и содержание Ознакомительных и Валидационных датасетов не являются полностью идентичными. Оргкомитет имеет право изменять количество Эпикризов Ознакомительного датасета.
- 1.10. **Обоснование корректности работы СППВР** - статистическая метрика, характеризующая степень достоверности результата СППВР.
- 1.11. **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.
- 1.12. **Полнота объяснения** - наличие в ответе СППВР следующих сущностей: предварительный диагноз (предположение о диагнозе в виде кода МКБ10), набор запросов исследований, заключительный клинический диагноз.
- 1.13. **Понятность объяснения** - по мнению Клинического специалиста: предварительный диагноз соответствует первичным данным клинического случая, набор запросов исследований соответствует предварительному диагнозу, заключительный клинический диагноз корректен и соответствует траектории и результатам исследований.
- 1.14. **Программный комплекс участника (ПКУ, Продукт, СППВР)** – 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.
- 1.15. 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.
- 1.16. **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.

- 1.17. **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.
- 1.18. **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.
- 1.19. **Точность постановки диагноза** - под точностью в рамках Конкурса понимается совокупность параметров чувствительности Se и специфичности Sp, согласно методике расчета в приложениях 1а, 1б, 1в.
- 1.20. **Судья** – член Судейской коллегии.
- 1.21. **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.
- 1.22. **Эпикриз** - совокупность **Анонимизированных** данных о пациенте, содержащая результаты врачебных осмотров, результаты лабораторных и инструментальных исследований, а также прочие данные о клиническом случае. Эпикризы собираются из действующих информационных систем медицинских учреждений и верифицируются практикующими врачами высшей категории. При этом полнота и качество данных в Эпикризах соответствуют реальной ситуации в современной медицинской практике. **Также, в Эпикризах могут отсутствовать морфологические исследования (и ряд других высокотехнологичных 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 или этапа Конкурса, 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 клинико-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 настоящим Техническим регламентом, что является необходимым условием для победы в Финальном этапе. **Преодоление Технологического барьера - это преодоление пороговых значений Чувствительности и Специфичности по нижней границе доверительного интервала по всем нозологиям.**

5.5. В качестве пороговых значений Технологического барьера, учитываемых Интегральным критерием, используется чувствительность и специфичность в рамках конкретной нозологии **с уровнем доверия 95%.**

Нозология	Рак легкого	Туберкулез	Бактериальная пневмония	Вирусная пневмония (вкл. SARS-CoV-2)	ХТЭЛГ	ХОБЛ
Чувствительность	83,000%	83,000%	82,000%	82,000%	81,000%	83,000%
Специфичность	87,000%	87,000%	83,000%	83,000%	88,000%	84,000%

5.6. В случае, если после прохождения Испытаний Финального этапа ни одна из участвующих в Испытаниях **СППВР не продемонстрировала превышение пороговых значений Чувствительности и Специфичности по нижней границе доверительного интервала** по всем нозологиям – **возможно только Частичное преодоление Технологического барьера. Команда может претендовать на Частичное преодоление Технологического барьера** в том случае, если в рамках прохождения Испытаний Финального этапа было зафиксировано превышение пороговых значений Чувствительности и Специфичности по нижней границе доверительного интервала для СППВР Команды хотя бы по одной нозологии. Условия присуждения Частичного преодоления Технологического барьера и ранжирования Команд описаны в Приложении 1в.

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 официальном сайте Конкурса - <https://ai4med.upgreat.one/>).

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, утвержденном требованиями Приложений 2,3,4) 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 J16 J18
Вирусная пневмония (включая SARS-CoV2)	J10 J12 U07
Хроническая обструктивная болезнь легких (ХОБЛ)	J44
Хроническая тромбоэмболическая легочная гипертензия (ХТЭЛГ)	I27 (Эпикризы с подтвержденной нозологией ХТЭЛГ соответствуют клинической картине МКБ 10: I27.8. Однако, в рамках унифицированного

	подхода к формату ответа СППВР используется только базовая часть кода - I27)
Здоровый человек (не болен ни одной из вышеперечисленных нозологий)	В эпикризах может присутствовать любой код МКБ 10, кроме указанных выше. Однако ответ СППВР Платформе должен соответствовать требованиям, указанным в Приложениях 2,3 (значение "another" в ключе "diagnosisMain")

- 6.6. Нозологии, представленные в п.6.5 являются равнозначными в рамках Конкурса и не ранжируются с точки зрения значимости/опасности в медицинской практике.
- 6.7. Для участия в Испытаниях Полуфинального и Финального этапов Команда должна запустить Сессию обмена данными с Платформой до момента старта Испытаний и по команде Судьи. В случае подключения к Платформе после начала Испытаний, метрика для пропущенных Командой Эпикризов аннулируется, диагноз считается поставленным неверно.
- 6.8. 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.
- 6.9. 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, **либо предоставляются для общего ознакомления Команд и не учитываются для расчета Интегрального критерия и в дальнейшем при ранжировании результатов Команд.** Даты проведения Тестовых испытаний сообщаются Командам не позднее чем за 5 (пять) рабочих дней до момента проведения.
- 6.10. 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.
- 6.11. 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.

- 6.12. На протяжении всех этапов Конкурса Оргкомитетом может предоставляться канал коммуникации для оперативного оповещения Команд о мероприятиях, процессах и требованиях Конкурса.
- 6.12.1. В качестве канала коммуникации могут использоваться инструменты мессенджеров или сред совместной работы.
- 6.12.2. Члены Команд обязаны иметь в канале коммуникации (если это предусмотрено способом коммуникации) имя пользователя (ник), соответствующий следующей структуре: **Название Команды_ФИО члена Команды полностью**.
- 6.13. 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:
- 6.13.1. В качестве даты новых Испытаний может быть назначен другой день (но не позднее чем 7 (семь) рабочих дней с момента отмены Испытаний), о чём Оргкомитет уведомляет отдельно.
- 6.13.2. Используется Запасной датасет, размер и наполнение которого определяется Оргкомитетом. Все предыдущие результаты Команд в рамках данного этапа могут быть аннулированы **по решению Судейской Коллегии и Технической Комиссии; в этом случае аннулирование производится для всех Команд одновременно**. Регламент действий в случае перезапуска определяется Оргкомитетом и доводится до сведения Участников (посредством официальной электронной почты) не позднее чем за 5 дней до даты новых Испытаний.
- 6.14. Использование общих программных модулей и установленных разработчиком Платформы способов взаимодействия с Платформой, является обязательным, поскольку обеспечивает равные условия для всех Участников Конкурса.
- 6.15. Участники Конкурса не имеют права самостоятельно вносить изменения в общие программные модули Платформы.
- 6.16. Оргкомитет имеет право приостановить ход Испытаний любого из этапов, например, в **случае возникновения угрозы** Массового технического сбоя и его оперативного **предотвращения**. 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 одних Испытаний не может превышать 2 часа (фиксируется Судьями). В случае необходимости большей продолжительности Оргкомитет принимает решение о переносе Испытаний.

7. Отборочный этап

- 7.1. Данный раздел дополняет п.5.2 Конкурсного задания Конкурса
- 7.2. Задачей Отборочного этапа является формирование списка Команд, готовых технически и организационно к процессу работы в условиях 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 настоящего Технического регламента).
- 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 по решению Оргкомитета после проведения Испытаний 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. Предположение о заключительном клиническом диагнозе по каждому из Эпикризов Валидационного датасета на основе неполных данных.
 - 8.4.2. Постановка Заключительного клинического диагноза по каждому из Эпикризов Валидационного датасета.
- 8.5. Результаты работы СППВР передаются на Платформу (в Систему проверки решений). Система проверки решений сравнивает результат работы СППВР Команд с эталонным результатом, предварительно сформированным

экспертной группой врачей высшей категории. По результатам сравнения рассчитываются метрики:

- 8.5.1. точности (чувствительность и специфичность) выявления каждой из нозологий основного диагноза;
 - 8.5.2. выявления осложнений и сопутствующих заболеваний. Выполняется Экспертной комиссией в случае равенства результатов по точности диагностики основного заболевания;
 - 8.5.3. ресурсоемкости (стоимости) постановки диагноза. Выполняется на Финальном этапе для Команд, допущенных к Финальному этапу.
- 8.6. Испытания Полуфинального этапа проводятся на основе Валидационного датасета, имеющего объем не менее 500 Эпикризов. Диагнозы в Валидационном датасете Полуфинального этапа отсутствуют.
- 8.7. Каждый элемент Валидационного датасета представлен в двух версиях - неполные данные (версия v3) и полные данные (версия v2). Состав версий может различаться в разных элементах датасета. В некоторых случаях версии Эпикризов v2 и v3 одного и того же клинического случая могут иметь минимальные отличия или не иметь отличий. Данные случаи считаются нормой в рамках условий проведения Конкурса.
- 8.8. Во время Испытаний Полуфинального этапа элементы Валидационного 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 и размещается на Сайте.
- 8.9. Испытания Полуфинального этапа проходят следующим образом:
- 8.9.1. Запуск Сессии обмена данными между СППВР и Платформой.
 - 8.9.2. Платформа публикует случайный Эпикриз в версии неполных данных (v3).
 - 8.9.3. 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.
 - 8.9.4. Платформа публикует полную версию текущего элемента Валидационного датасета (v2).
 - 8.9.5. СППВР Команды анализирует Эпикриз, и на основе информации в нем возвращает:
 - код МКБ 10 Основного заболевания (согласно п.6.5), либо отсутствие заболевания в рамках нозологий Конкурса;
 - коды МКБ 10 сопутствующих заболеваний и Осложнений;

На данную операцию отводится 30 секунд.

- 8.10. Формат ответа, необходимого к передаче на Платформу со стороны СППВР, подробно описан в Приложении 3.
- 8.11. Если в процессе Сессий обмена СППВР Команды не возвращает ответ на Платформу в течение 30 секунд после публикации Эпикриза - текущий Эпикриз аннулируется для данной Команды. В этом случае диагноз считается поставленным неверно.
- 8.12. Если аннулирование 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.
- 8.13. По окончании Испытаний Сессии обмена между СППВР и Платформой закрываются, все данные, полученные от СППВР, сохраняются на Платформе, формируя закрытый датасет с полуфинальными Совокупностями данных Команд. С этого момента стартует процесс оценивания результатов Команд.
- 8.14. После предварительной ревизии Совокупностей данных запускается процесс расчета Интегрального критерия Команд.
- 8.15. По итогам Полуфинального этапа строится лидерборд — список Команд, ранжированный по убыванию обобщенного качества СППВР S^k (см. Приложение 1а, 1б). Аннулированные Совокупности данных не учитываются, а соответствующие им СППВР не участвуют в лидерборде.
- 8.16. К финальному этапу допускается не более 4 (четырёх) Команд с наивысшим значением обобщенного качества СППВР S^k .
- 8.17. В случае, если СППВР Команды по результатам Полуфинального этапа соответствует критериям преодоления Технологического барьера или критериям Частичного преодоления Технологического барьера - это ни при каких обстоятельствах не может является условием победы Команды в Конкурсе. Победа в Конкурсе и преодоление Командой Технологического барьера или Частичное преодоление Технологического барьера может фиксироваться только по результатам Испытаний Финального этапа.
- 8.18. Если по результатам Полуфинального этапа не может быть отобрано 4 (четыре) Команды для участия в Финальном этапе, то Полуфинальный этап может быть проведен повторно по решению Оргкомитета.
- 8.19. В процессе проведения Полуфинального этапа Командам предоставляется доступ к информации медицинских баз данных (клинические рекомендации). Доступ предоставляется через базу знаний Платформы.

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 членов Команды на **Конкурсной площадке**.
- 9.3. Испытания Финального этапа проводятся на основе Валидационного датасета Финального этапа, имеющего объем не менее 100 (ста) Эпикризов.
- 9.4. Каждый Эпикриз Валидационного датасета Финального этапа представлен следующими версиями файлов.

Версия файла	Содержание	Доступность для СППВР
v1	Полный Эпикриз, содержащий все доступные данные по клиническому случаю, включая диагнозы и высокотехнологические исследования.	Ознакомительный датасет.
v3F	Эпикриз, содержащий начальные данные по клиническому случаю (без инструментальных и лабораторных исследований).	Ознакомительный датасет, Валидационный датасет
v32	Множество файлов, относящихся к одному эпикризу. Каждый файл содержит информацию об одном исследовании (v32I - инструментальном; v32L - лабораторном)	Ознакомительный датасет, Валидационный датасет

- 9.5. Базовый нейминг файлов в рамках каждого эпикриза представлен следующей структурой:

AIDoc_нозология_№эпикриза_версия_№исследования в эпикризе_код исследования.xml

Пример: AIDoc_Tuberkulez_112_v32L_2L_B03.016.002.xml

В рамках Валидационного датасета базовый нейминг заменяется для СППВР условными номерами файлов для обезличивания принадлежности к нозологиям.

- 9.6. Концепция Испытаний Финального этапа заключается в **пошаговой** постановке диагноза в условиях неполных или недостаточных анонимизированных данных о пациенте. Изначально Команда получает первичный набор анонимизированных данных о пациенте (v3F), после чего должна, **выдвинув предположение о диагнозе**, запрашивать через Платформу необходимые исследования (v32) и в завершении цикла поставить **заключительный клинический диагноз**.
- 9.7. В рамках данного этапа Продуктам Команд необходимо:

- 9.7.1. Определить заключительный клинический диагноз по всем Эпикризам Валидационного датасета, предназначенных для Финального этапа Конкурса.
- 9.7.2. Продемонстрировать Клиническому специалисту способность к прогнозированию диагноза по неполным данным, определению необходимой последовательности исследований для подтверждения диагноза, анализу результатов исследований и формированию заключительного клинического диагноза.
- 9.8. Критерии оценивания СППВР на Финальном этапе Конкурса представлены в Приложении 1в. Наиболее значимой метрикой является точность постановки основного диагноза, выраженная в виде совокупности (обобщенное качество S^*) Чувствительности и Специфичности по нозологиям. В Интегральном критерии метрика точности является главенствующей над остальными метриками (ресурсоемкость, точность определения осложнений и сопутствующих заболеваний, обоснованность и т.д.).
- 9.9. Клинические специалисты, в процессе парной работы с СППВР, оценивают Интерпретируемость, Полноту объяснения, Понятность объяснения СППВР. Обоснование корректности работы СППВР выполняется на основе математического расчета (см. Приложение 1в).
- 9.10. Испытания Финального этапа проходят в две стадии.
 - 9.10.1. Стадия 1: Пошаговая постановка диагнозов в режиме ограниченного времени и в автоматическом режиме. Клинический специалист не участвует.
 - 9.10.1.1. ШАГ 1. Платформа отправляет всем СППВР случайный Эпикриз версии V3F из Валидационного датасета Финального этапа.
 - 9.10.1.2. ШАГ 2. СППВР обрабатывает файл, и на основе информации в нем возвращает на Платформу предположение об основном диагнозе (в виде кода МКБ10).
 - 9.10.1.3. ШАГ 3. СППВР выполняет одно из следующих действий:
 - 9.10.1.3.1. ШАГ 3.1 Возвращает значение заключительного клинического диагноза на Платформу. После этого действия работа с текущим Эпикризом завершается. Производить дальнейший запрос исследований по данному Эпикризу - невозможно.
 - 9.10.1.3.2. ШАГ 3.2 Запрашивает дополнительные исследования (инструментальные или лабораторные) по коду запроса (возможен только один код в одном запросе) исследования из таблицы исследований (см. Приложение 4). В

зависимости от наличия исследования в Эпикризе
Платформа:

- 9.10.1.3.2.1. ШАГ 3.2.1 Вернет один или несколько файлов v32, содержащих информацию по запрошенному исследованию или группе исследований. Стоимость исследования (исследований) добавляется к стоимости ранее выполненных исследований;
- 9.10.1.3.2.2. ШАГ 3.2.2 Вернет отчет об отсутствии данного исследования в составе Эпикриза. Стоимость исследования (исследований) добавляется к стоимости ранее выполненных исследований.
- 9.10.1.3.3. ШАГ 3.3 СППВР анализирует полученные данные и переходит к действиям шага 3.
- 9.10.1.4. Действия шага 3 повторяются циклически и завершаются в случаях:
 - СППВР поставлен заключительный клинический диагноз по текущему эпикризу;
 - процесс постановки диагноза прерван со стороны СППВР ввиду невозможности поставить диагноз;
 - истекло время, отведенное на постановку диагноза (1(одна) минута).
- 9.10.1.5. Переход на шаг 1.
- 9.10.2. На каждой итерации (ШАГ 3) Платформа сохраняет данные о процессе Испытаний (запрошенное исследование, время запроса, статус ответа), протоколируя траекторию постановки диагноза СППВР.
- 9.10.3. На шаге 3.2 (запрос дополнительного исследования) разрешается производить запрос только на одно исследование. Указание в одном запросе нескольких исследований - недопустимо.
- 9.10.4. В случае, если в Эпикризе содержится несколько однотипных исследований (разные исследования, соответствующие одному коду в рамках Конкурса; одинаковые исследования, выполненные в разное время) Платформа вернет несколько файлов v32.
- 9.10.5. В случае, если СППВР на шаге 2 не предоставляет на Платформу предположение об основном диагнозе - считается, что по данному Эпикризу СППВР не продемонстрировала Интерпретируемость, Полноту объяснений и Понятность объяснений.

9.10.6. После обработки всех файлов Валидационного датасета Полуфинального этапа сессия обмена с Платформой завершается, стадия 1 считается завершенной.

9.10.7. Экспертной группой врачей на основе анализа результатов СППВР формируется Валидационный датасет стадии 2. Данный датасет формируется из эпикризов, использующихся на стадии 1. Объем Валидационного датасета стадии 2 - не менее 16 и не более 32 эпикризов.

9.10.8. Стадия 2: Работа в паре с Клиническим специалистом (интерпретирование)

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

9.10.8.2. На стадии 2 Клинический специалист работает в паре с СППВР (информацией о результатах диагностики, проведенной СППВР по клиническому случаю). Непосредственное взаимодействие СППВР и Клинического специалиста не осуществляется.

9.10.8.3. Платформа воспроизводит работу СППВР для Клинического специалиста по траектории, сохраненной на стадии 1. При этом Платформа может одновременно воспроизводить работу одной СППВР для нескольких клинических специалистов.

9.10.8.4. Работа Клинических специалистов проводится в несколько раундов. В каждом раунде восемь Клинических специалистов одновременно работают с СППВР Команд финалистов (четыре команды). В одном раунде всем Клиническим специалистам предоставляется доступ к одному и тому же Эпикризу. Эпикриз заменяется в каждом раунде. Общая структура взаимодействия пар в раунде:



9.10.8.5. Пары (Клинический специалист - СППВР) формируются случайным образом (жеребьевкой). Жеребьевка проводится Судейской коллегией на Конкурсной площадке при участии представителей Команд. После жеребьевки СППВР и Клинические специалисты получают условные имена (C1,C2,C3,C4), (КС1,КС2,КС3,КС4,КС5,КС6,КС7,КС8), которые заносятся на Платформу. Жеребьевка проводится один раз, после чего в последующих раундах происходит изменение пар по следующей схеме:

	СППВР	КС	№ Эпикриза		СППВР	КС	№ Эпикриза

РАУНД 1	C1	KC1	KC5	1		РАУНД 3	C1	KC3	KC7	3
	C2	KC2	KC6	1			C2	KC4	KC8	3
	C3	KC3	KC7	1			C3	KC1	KC5	3
	C4	KC4	KC8	1			C4	KC2	KC6	3
РАУНД 2	C1	KC2	KC6	2		РАУНД 4	C1	KC4	KC8	4
	C2	KC3	KC7	2			C2	KC1	KC5	4
	C3	KC4	KC8	2			C3	KC2	KC6	4
	C4	KC1	KC5	2			C4	KC3	KC7	4

- 9.10.8.6. Количество раундов на стадии 2 определяется Оргкомитетом и может составлять от 16 до 32 раундов.
- 9.10.8.7. В процессе стадии 2 Клинические специалисты формируют заключение о Понятности объяснения и Полноте объяснения СППВР. На основе заключения Клинических специалистов формируется оценка интерпретируемости СППВР.
- 9.11. После окончания Испытаний Совокупности данных СППВР и комментарии Клинических специалистов сохраняются на Платформе, формируя закрытый датасет результатов Финального этапа.
- 9.12. Платформа производит расчет Интегрального критерия по методике, описанной в Приложении 1а, 1в. В расчете результатов Финального этапа учитываются результаты СППВР на Испытаниях Полуфинального этапа (точность и ресурсоемкость).
- 9.13. Испытания Финального этапа проводятся очно на территории **Конкурсной площадки**. В процессе проведения Испытаний Финального этапа на Конкурсной площадке должны присутствовать не менее двух представителей Команды.
- 9.14. Процесс Испытаний Финального этапа 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. **Конкурсная площадка**

- 10.1. Стадии 1 и 2 Испытания Финального этапа проходят на Конкурсной площадке с очным присутствием представителей Команд. В состав Конкурсной площадки включены следующие зоны:
- 10.1.1. Техническая зона – зона подготовки Команд к проведению Испытаний. Каждой Команде в рамках технической зоны предоставляется рабочее место содержащее в минимальном оснащении: точку подключения к сети 220 В, стол и стулья для размещения членов Команды, точка подключения к Интернет (проводным или беспроводным способом).

Оргкомитет может изменять количественный и качественный состав технического оснащения Технической зоны для Команд.

- 10.1.2. Зона испытаний - специально подготовленная часть Конкурсной площадки, используемая для проведения стадии 2.
- 10.1.3. Зона размещения Оргкомитета, Экспертной комиссии и Судейской коллегии.
- 10.2. В процессе Испытаний и всех подготовительных этапов члены Судейской коллегии и Технической комиссии, а также представители Оргкомитета имеют право находиться во всех зонах Конкурсной площадки. Возможность посещения членами Команд различных зон Конкурсной площадки определяются на Финальном этапе до начала Испытаний и доводится до Команд на установочном брифинге.
- 10.3. Для осуществления процессов запуска СППВР, и других действий в процессе Испытаний, Команда должна использовать собственное компьютерное оборудование (ноутбук, планшет и т.д.). Оргкомитет не предоставляет Командам компьютерное оборудование в процессе Испытаний Конкурса.
- 10.4. Требования внутреннего распорядка на Конкурсной площадке, а также иные правила, включая правила безопасности, предоставляются Командам на Финальном этапе (не позднее 10 дней до начала Испытаний Финального этапа) и должны соблюдаться Командами в течение всего времени проведения Испытаний.

11. Судейство Конкурса

- 11.1. Судейская коллегия проверяет автоматически формируемые заключения о прохождении Командой алгоритмической квалификации и готовности к Испытаниям Полуфинального и Финального этапов.
- 11.2. Результаты Испытаний Полуфинального и Финального этапов верифицируются Технической комиссией и передаются на утверждение Оргкомитету.
- 11.3. Техническая комиссия проверяет корректность полученных от Команды данных, выявляет условия и причины аннулирований на стороне Платформы, или причины Массового технического сбоя, если таковые имеются. В данном случае Судейской коллегией составляется соответствующий протокол и передается в Оргкомитет.
- 11.4. По окончании Испытаний Финального этапа и процедур расчета интегрального критерия Судейская коллегия проверяет автоматически формируемые отчеты о результатах Команд и условия преодоления технологического барьера.
- 11.5. Процесс определения победителей подробно описан в п.6 и п.7 Конкурсного задания Конкурса.
- 11.6. Оргкомитет выносит итоговые решения по результатам Полуфинального и Финального этапов, которые передаются на рассмотрение и утверждение Жюри.

12. Ограничения и дисквалификации

- 12.1. Участники и члены Команд Конкурса обязаны следовать требованиям Конкурсного Задания и Технического регламента Конкурса.
- 12.2. Все решения по дисквалификации Команды и аннулированию результатов принимаются Оргкомитетом и вступают в силу незамедлительно.
- 12.3. 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 или членов Команд, воздерживаться от оскорблений в адрес других Участников или членов Команд, Оргкомитета и партнеров Конкурса.
- 12.4. Участнику и Командам запрещается осуществлять экспертную разметку Тестового или Валидационного датасета вручную, с помощью краудсорсинга или любыми иными способами с привлечением труда людей.
- 12.5. Запрещаются действия направленные на попытку определить персональные данные субъекта Эпикриза (пациента) путем анализа сходства результатов клинических исследований или любым другим путем. Подобные действия будут расценены как нарушение Участником законодательства о персональных данных.
- 12.6. 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.).
- 12.7. 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.
- 12.8. На всех этапах конкурса при работе с Платформой запрещаются любые обращения к Платформе с частотой выше 10 раз в секунду.
- 12.9. 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.
- 12.10. 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 о соответствии решений целям конкурса Команда имеет право запросить официальную позицию Оргкомитета.

- 12.11. Оргкомитет имеет право дисквалифицировать Команду в случае отказа от выполнения требований внутреннего распорядка и требований безопасности, а также отказа от выполнения требований нормативных документов, относящихся к проведению Испытаний на Конкурсной площадке.
- 12.12. 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).

13. **Особые случаи и обеспечение безопасности Конкурса**

- 13.1. Если в процессе Испытаний Команда отправляет на Платформу несколько ответов на один запрос, то алгоритмы Платформы учитывают только последний **ответ**.
- 13.2. 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).
- 13.3. 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.
- 13.4. Деятельность Участников в рамках Конкурса должна соответствовать действующим на территории Российской Федерации природоохранным нормам, требованиям техники безопасности.
- 13.5. На финальном этапе Конкурса Оргкомитет может предоставить Командам правила по технике безопасности, с которыми все члены Команд должны ознакомиться под подпись и соблюдать.

14. **Дополнительные условия**

- 14.1. Оргкомитет оставляет за собой право обновлять Ознакомительный и Валидационный датасеты, а также исходные коды общих программных модулей Платформы, но не позднее, чем за 10 (десять) календарных дней до Испытаний Финального этапа. В случае обновления Оргкомитет оповещает всех Участников путем публикации на сайте Платформы, сообщая цель и суть произведенных модификаций.

- 14.2. В случае обнаружения ошибок в математическом, алгоритмическом обеспечении, датасетах, общих программных модулях Платформы Оргкомитет оставляет за собой право обновлять Ознакомительный и Валидационный датасеты, а также исходные коды общих программных модулей Платформы в любой момент, кроме времени непосредственно Испытаний. В данном случае информирование Участников может не проводиться.
- 14.3. Оргкомитет имеет право проверить решение Команды очно или провести дополнительные процедуры проверки, в случае если у Судейской коллегии или Технической комиссии возникнут основания считать Совокупность данных какой-либо Команды не соответствующим целям и задачам Конкурса.
- 14.4. В случае, если Участник Конкурса желает прекратить участие своей Команды в Конкурсе на любом этапе - необходимо отправить заявление от лица Участника и с официальной почты Участника в Оргкомитет. В данном случае порядок действий по прекращению участия Команды в Конкурсе определяется Оргкомитетом.
- 14.5. Оргкомитет имеет право публиковать на Сайте конкурса результаты Команд (в том числе полные Совокупности данных) после Испытаний Финального этапа с целью обеспечения прозрачности в определении победителя и призёров.
- 14.6. 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.
- 14.7. В соответствии с п. 8.2. Конкурсного задания данные, сформированные СППВР во время Испытаний (**Совокупности данных**), 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 1а

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) представляются матрицей ошибок (Рис.1) в которой A , B , C и D — количественные результаты диагностики системой по выборке пациентов N , соответствующих каждому результату диагностической классификации. В качестве примера матрицы ошибок на рисунке 1 представлены результаты диагностического теста для i -ой нозологии, по которой качество диагностики оценивается по выборке из 120 эпикризов с распределением по статусам 'болен' – 75 'не_болен' - 45.

Нозология		i		$i=[1..6]$
Результаты теста по нозологии 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)

Рисунок 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 матрицу ошибок, как классификационную таблицу, формулы для расчета чувствительности и специфичности по выборке результатов диагностических тестов выглядят следующим образом:

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

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

Ввиду вероятностного характера специфичности и чувствительности, оцениваются лишь их статистики, которые в дальнейшем обозначим как:

$$Se = \frac{T1}{n1}, Sp = \frac{F0}{n0}, \quad (4)$$

где:

$T1$ — количество больных диагностированных верно (TP - *TruePositive*) ;

$n1$ — общее количество больных в исследуемой выборке N ;

$F0$ — количество 'не больных' диагностированных как 'не больные' (TN - *TrueNegative*) ;

$n0$ — общее количество 'не больных' 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 технологического барьера по каждой i -ой нозологии $i=[1; 6]$, т.е. выполнение неравенства:

$$Se_i^j > Se_i^{TB}; Sp_i^j > Sp_i^{TB}, \quad (5)$$

где Se_i^{TB} и Sp_i^{TB} - чувствительность и специфичность по конкретной нозологии.

Так как Se_i^j , Sp_i^j - вероятностные величины, то преодоление технологического барьера по ним оценивается по нижним границам их доверительных интервалов, с вероятностью не ниже 95%. Методика их расчета содержится в п.2.

При получении одинаковых статистических оценок Se , Sp для диагностики основного заболевания Командами Конкурса, 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 трех знаков после запятой. Округление 4-го знака осуществляется по правилам математики от 0,0005 в большую сторону.

Формирование матриц ошибок для нозологий Конкурса осуществляется по схеме отображенной на Рисунке 2.

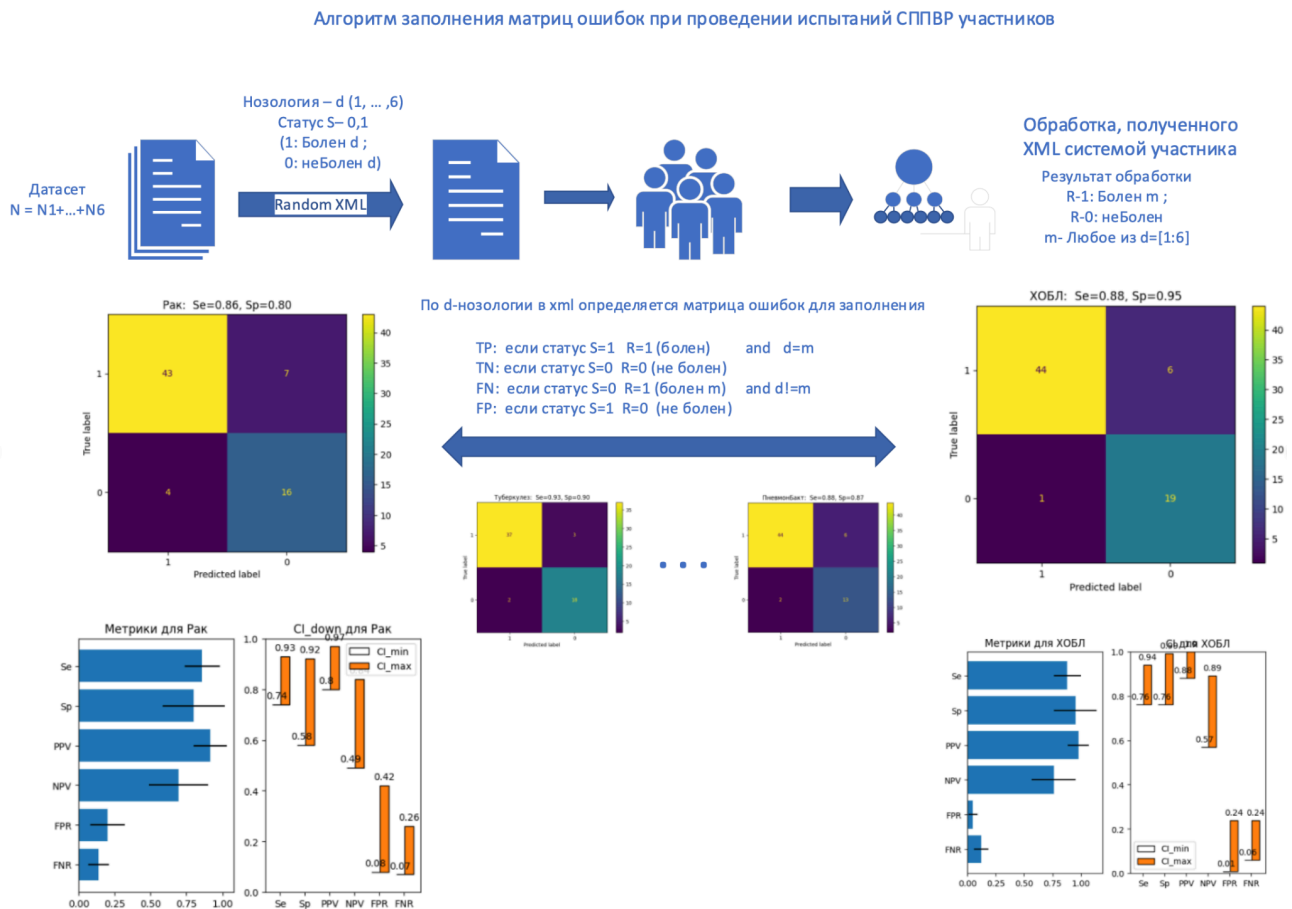


Рисунок 2 - Формирование матриц ошибок для нозологий при испытаниях СППВР

2. Доверительные интервалы для оценки показателей чувствительности и специфичности

Расчетные значения чувствительности и специфичности (4) носят вероятностный характер, поэтому для их статистик 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 заболевания, интервал в котором истинное значение специфичности диагностического теста будет находиться с вероятностью не менее 0,95.

Доверительные интервалы для чувствительности определяются по (6)

$$CI = \frac{Se + \frac{z_{1-\alpha/2}^2}{2n1} \pm z_{1-\alpha/2} \sqrt{\frac{Se(1-Se) + \frac{z_{1-\alpha/2}^2}{4n1}}{n1}}}{1 + \frac{z_{1-\alpha/2}^2}{n1}}, \quad (6)$$

Доверительные интервалы для специфичности определяются по (7)

$$CI = \frac{Sp + \frac{z_{1-\alpha/2}^2}{2n0} \pm z_{1-\alpha/2} \sqrt{\frac{Sp(1-Sp) + \frac{z_{1-\alpha/2}^2}{4n0}}{n0}}}{1 + \frac{z_{1-\alpha/2}^2}{n0}}, \quad (7)$$

где z - значение γ -квантиля нормального распределения.

В последующих расчетах из (6) и (7) используются только их нижние границы. Поэтому z -значение γ -квантиля нормального распределения, $\gamma = 1-\alpha/2$ для двустороннего интервала $\alpha = 0,05$ $\gamma = 0,975$, заменяем на значение для одностороннего интервала $\gamma = 1-\alpha$, где $\alpha = 0,05$ $\gamma = 0,95$. Для односторонней границы доверительного интервала значение $z_{1-\alpha}$ составит 1.64. Окончательные формулы для расчета нижних границ доверительных интервалов чувствительности и специфичности (8) представлены ниже:

$$CI_{min-Se} = \frac{Se + \frac{z_{1-\alpha}^2}{2n1} - z_{1-\alpha} \sqrt{\frac{Se(1-Se) + \frac{z_{1-\alpha}^2}{4n1}}{n1}}}{1 + \frac{z_{1-\alpha}^2}{n1}},$$

$$CI_{min-Sp} = \frac{Sp + \frac{z_{1-\alpha}^2}{2n0} - z_{1-\alpha} \sqrt{\frac{Sp(1-Sp) + \frac{z_{1-\alpha}^2}{4n0}}{n0}}}{1 + \frac{z_{1-\alpha}^2}{n0}}. \quad (8)$$

Таким образом, условием преодоления технологических барьеров с уровнем вероятности 95%, для Участников конкурса является условие преодоления ими нижних границ односторонних доверительных интервалов для чувствительности и специфичности по каждой нозологии, т.е. выполнение неравенств:

$$CI_{min_}Se_i^j > Se_i^{TB}; CI_{min_}Sp_i^j > Sp_i^{TB}, \quad (9)$$

где:

Se_i^{TB} и Sp_i^{TB} - технологические барьеры для чувствительность и специфичность по i -нозологии конкурса, $i = [1:6]$ - нозологии конкурса, j - индекс Участника конкурса.

3. Количество итераций и учет ресурсоемкости (стоимости) в Интегральном критерии (Финальный этап)

Для учета количества потраченных итераций и ресурсоемкости испытания СППВР Участников осуществляются посредством последовательной подачи на вход модели Участников xml-файлов отличающихся полнотой лабораторных и инструментальных исследований содержащихся в них. Файл **xml_v3** может содержать ограниченный перечень лабораторных и инструментальных исследований по сравнению с версией **xml_v2**. Также, используется множество отдельных исследований, которые может запрашивать СППВР (совокупность **v32**).

В соответствующих <entry> секциях файлов xml_v3, v2, v32 содержится уникальный идентификатор исследования и код справочника. Например для инструментального исследования - "Компьютерная томография груди с внутривенным контрастированием":

```
code code="7001902"
codeSystem="1.2.643.5.1.13.13.11.1471"
codeSystemName="Федеральный справочник инструментальных диагностических исследований" ...
```

По коду "7001902" исследования определяется его наименование и код НМУ. По наименованию медицинской услуги платформа конкурса получает ресурсоемкость из Приложения 1 Номенклатуры работ и услуг в здравоохранении утвержденных Минздравсоцразвития РФ 12.07.2004, либо другими нормативными документами Минздрава России и Федерального Фонда ОМС.

Ресурсоемкость услуги считается, как сумма значений УЕТ врача и УЕТ м/с .

Если в Приложении 1 отсутствует услуга представленная в xml-документе, то для этого исследования принимается нулевая ресурсоемкость.

Таким образом, по количеству запрошенных исследований Участником для постановки диагноза, оценивается ресурсоемкость (стоимость, трудоемкость), которую он потратил на диагностику.

Обозначим:

- d – нозологии конкурса $d = 1, \dots, 6$;
- L_d – перечень лабораторных исследований в датасете для d -ой нозологии;
- C_d^L – ресурсоемкость (стоимость) лабораторного исследования для d -ой нозологии;
- I_d – перечень инструментальных исследований, присутствующие в датасете для d -ой нозологии;
- C_d^I – ресурсоемкость (стоимость) инструментального исследования для d -ой нозологии;
- C - оценка ресурсоемкости (стоимости).

$$C = C_d^L + C_d^I = \sum_n \sum_d C_d^n(I_d) + \sum_n \sum_d C_d^n(L_d). \quad (10)$$

Накапливая таким образом стоимость диагностики, получим оценку ресурсоемкости (стоимости) для каждого Участника

$$C += C_d^L + C_d^I. \quad (11)$$

Примечание

На Полуфинальном этапе сбор информации о ресурсоемкости определяется по версиям v3 (минимальная ресурсоемкость $C_{\min}$) и v2 (максимальная ресурсоемкость $C_{\max}$). $C_{\min}$, $C_{\max}$ - суммарные значения ресурсоемкости версий v3 и v2 соответственно. Суммарная ресурсоемкость определяется как сумма значений ресурсоемкости всех лабораторных, инструментальных и иных исследований в Эпикризе.

Приложение 16.

Алгоритм определения победителей Полуфинального этапа

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

K – количество команд участвующих в конкурсе

Каждый экземпляр **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.

Для каждого Участника конкурса $k = [1;K]$ выполняем последовательно алгоритм, изображенный на рисунке 3.

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 один из результатов: ['Рак', 'Туберкулез', 'Бактериальная пневмония', 'Вирусная пневмония', 'ХОБЛ', 'ХТЭЛГ', 'не болен'] .

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.

$$TP += TP_i; \quad TN += TN_i; \quad FH += FN_i; \quad FP += FP_i. \quad (12)$$

Полученная ресурсоемкость добавляется к накапливаемой ресурсоемкости диагностики СППВР Команды.

$$C_k += C_i. \quad (13)$$

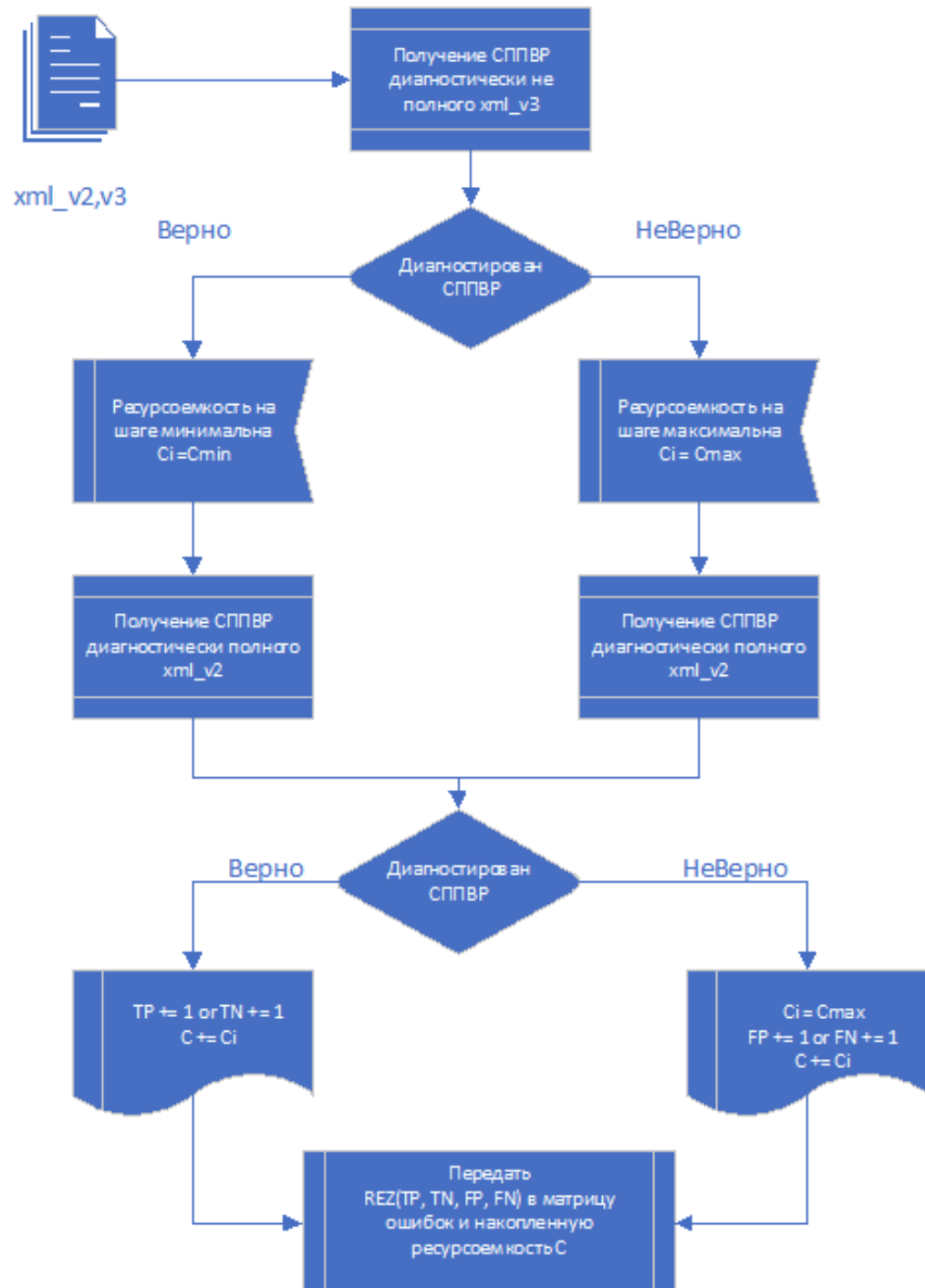


Рис. 3 Алгоритм оценки точности и ресурсоемкости диагностики заболевания для нозологии Конкурса

Пример сценариев заполнения матрицы ошибок на примере нозологии “Рак легкого”:

Все результаты по данной таблице идут в матрицу "Рак легкого"			
№	Подано от Платформы	Ответ СППВР	Действие системы
1	V3 Рак легкого	МКБ-10 Рак легкого	Запись свойства "Минимальная ресурсоемкость"
	V2 Рак легкого	МКБ-10 Рак легкого	Истинно положительный ответ с минимальной ресурсоемкостью
2	V3 Рак легкого	МКБ-10 Рак легкого	Запись свойства "Минимальная ресурсоемкость"
	V2 Рак легкого	МКБ-10 любого другого заболевания/Не болен ни одной из нозологий конкурса	Ложноотрицательный ответ с максимальной ресурсоемкостью
3	V3 Рак легкого	МКБ-10 любого другого заболевания/Не болен ни одной из нозологий конкурса	Запись свойства "Максимальная ресурсоемкость"
	V2 Рак легкого	МКБ-10 Рак легкого	Истинно положительный ответ с максимальной ресурсоемкостью
4	V3 Рак легкого	МКБ-10 любого другого заболевания/Не болен ни одной из нозологий конкурса	Запись свойства "Максимальная ресурсоемкость"
	V2 Рак легкого	МКБ-10 любого другого заболевания/Не болен ни одной из нозологий конкурса	Ложноотрицательный ответ с максимальной ресурсоемкостью
5	V3 Рак легкого	Нет ответа, либо ответ просрочен	Запись свойства "Максимальная ресурсоемкость"
	V2 Рак легкого	МКБ-10 Рак легкого	Истинно положительный ответ с максимальной ресурсоемкостью
6	V3 Рак легкого	МКБ-10 Рак легкого	Запись свойства "Минимальная ресурсоемкость"
	V2 Рак легкого	Нет ответа, либо ответ просрочен	Ложноотрицательный ответ с максимальной ресурсоемкостью
7	V3 Рак легкого	Нет ответа, либо ответ просрочен	Запись свойства "Максимальная ресурсоемкость"
	V2 Рак легкого	МКБ-10 любого другого заболевания/Не болен ни одной из нозологий конкурса	Ложноотрицательный ответ с максимальной ресурсоемкостью
8	V3 Рак легкого	МКБ-10 любого другого заболевания/Не болен ни одной из нозологий конкурса	Запись свойства "Максимальная ресурсоемкость"
	V2 Рак легкого	Нет ответа, либо ответ просрочен	Ложноотрицательный ответ с максимальной ресурсоемкостью
9	V3 Рак легкого	Нет ответа, либо ответ просрочен	Запись свойства "Максимальная ресурсоемкость"
	V2 Рак легкого	Нет ответа, либо ответ просрочен	Ложноотрицательный ответ с максимальной ресурсоемкостью

10	V3 Не рак легкого	Не болен ни одной из нозологий конкурса	Запись свойства "Минимальная ресурсоемкость"
	V2 Не рак легкого	Не болен ни одной из нозологий конкурса	Истинно отрицательный ответ с минимальной ресурсоемкостью
11	V3 Не рак легкого	Не болен ни одной из нозологий конкурса	Запись свойства "Минимальная ресурсоемкость"
	V2 Не рак легкого	МКБ-10 Любого заболевания	Ложноположительный ответ с максимальной ресурсоемкостью
12	V3 Не рак легкого	МКБ-10 Любого заболевания	Запись свойства "Максимальная ресурсоемкость"
	V2 Не рак легкого	Не болен ни одной из нозологий конкурса	Истинно отрицательный ответ с максимальной ресурсоемкостью
13	V3 Не рак легкого	МКБ-10 Любого заболевания	Запись свойства "Максимальная ресурсоемкость"
	V2 Не рак легкого	МКБ-10 Любого заболевания	Ложноположительный ответ с максимальной ресурсоемкостью
14	V3 Не рак легкого	Нет ответа, либо ответ просрочен	Запись свойства "Максимальная ресурсоемкость"
	V2 Не рак легкого	Не болен ни одной из нозологий конкурса	Истинно отрицательный ответ с максимальной ресурсоемкостью
15	V3 Не рак легкого	Не болен ни одной из нозологий конкурса	Запись свойства "Минимальная ресурсоемкость"
	V2 Не рак легкого	Нет ответа, либо ответ просрочен	Ложноположительный ответ с максимальной ресурсоемкостью
16	V3 Не рак легкого	Нет ответа, либо ответ просрочен	Запись свойства "Максимальная ресурсоемкость"
	V2 Не рак легкого	МКБ-10 Любого заболевания	Ложноположительный ответ с максимальной ресурсоемкостью
17	V3 Не рак легкого	МКБ-10 Любого заболевания	Запись свойства "Максимальная ресурсоемкость"
	V2 Не рак легкого	Нет ответа, либо ответ просрочен	Ложноположительный ответ с максимальной ресурсоемкостью
18	V3 Не рак легкого	Нет ответа, либо ответ просрочен	Запись свойства "Максимальная ресурсоемкость"
	V2 Не рак легкого	Нет ответа, либо ответ просрочен	Ложноположительный ответ с максимальной ресурсоемкостью

- После заполнения матриц ошибок для каждой нозологии $d = [1:6]$ рассчитываем показатели чувствительности и специфичности по формулам (2, 3);
- Для каждой метрики и по каждой нозологии рассчитываем нижние границы односторонних доверительных интервалов $CI_{min}Se$; $CI_{max}Sp$ по формулам (8). За

уровень значимости принимаем 95 % вероятность попадания полученных статистик для каждой метрики в границы интервала

$$\begin{aligned} Se(P = 95\%) &= [CI_{min}Se; 1], \\ Sp(P = 95\%) &= [CI_{min}Sp; 1]. \end{aligned} \quad (13)$$

4. Результаты расчетов каждой метрики d-ой нозологии сохраняем в таблице для нозологии d.

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

N п.п.	Метрика	Значение	Доверительный интервал	Стоимость ресурсоемкость (Усл.ед)
			CI_min	
1	Se (TPR)	92,00%	83,73%	7,968
2	Sp (TNR)	88,89%	76,66%	

5. Если показатели специфичности и чувствительности равны, рассчитываем Se по осложнениям вызванным основным заболеванием и ранжируем по Se без учета специфичности. (Процедура выполняется Технической комиссией).

6. Если показатели чувствительности для осложнений равны, рассчитываем Se по сопутствующим заболеваниям и ранжируем по Se без учета специфичности. (Процедура выполняется Технической комиссией).

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

Организаторы конкурса вправе установить или не устанавливать для чувствительности и специфичности критерии прохождения Участника в финальную часть конкурса $Se^{Kp}_d; Sp^{Kp}_d$ ¹.

Если такие критерии установлены, то Участники показавшие значения ниже этих критериев из Таблицы 2 удаляются и в последующем ранжировании не участвуют.

¹ Критерии прохождения в Финальный этап конкурса организаторами могут быть рассчитаны на основании допустимого уровня ошибок СППВР Участника, при которых обеспечивается преодоление технологических барьеров по нозологиям конкурса

Таблица 2 - Результаты испытаний СППВР в полуфинале Конкурса

Нозология Конкурса						
	Туберкулез	Пневмония бак	Пневмония вир	Рак легких	ХОБЛ	ХТЭЛГ
Критерии прохождения в финальную часть конкурса по чувствительности и специфичности						
SeS р	$Se^{Kp}_1; Sp^{Kp}_1$	...	...	...	...	$Se^{Kp}_6; Sp^{Kp}_6$
Результаты испытаний Участников						
1	$Se^I_1; Sp^I_1$ $CI_{min} Se^I_1;$ $CI_{min} Sp^I_1;$	...	...	...	...	$Se^I_D; Sp^I_D$ $CI_{min} Se^I_D;$ $CI_{min} Sp^I_D;$
	...	...	...	...	...	...
k	$Se^k_1; Sp^k_1$ $CI_{min} Se^k_1;$ $CI_{min} Sp^k_1;$	...	...	...	...	$Se^k_D; Sp^k_D$ $CI_{min} Se^k_D;$ $CI_{min} Sp^k_D;$
...	...	...	...	...	...	...
K	$Se^K_1; Sp^K_1$ $CI_{min} Se^K_1;$ $CI_{min} Sp^K_1;$	...	...	...	...	$Se^K_{ND}; Sp^K_D$ $CI_{min} Se^K_D;$ $CI_{min} Sp^K_D;$

Ранжирование, для отбора 4 (четырех) Команд в Финальный этап

Для каждого Участника считается геометрические средние по чувствительности и специфичности, показанные ими по D-нозологиям

$$Se^k_{cp} = \sqrt[6]{\prod_{d=1}^6 Se^k_d}, \quad (14)$$

$$Sp^k_{cp} = \sqrt[6]{\prod_{d=1}^6 Sp^k_d}.$$

Для последующей оценки вводится понятие обобщенного качества СППВР Участника S^k , которое определяется как геометрическое среднее показанных чувствительности и специфичности:

$$S^k = \sqrt[2]{Se^k_{cp} * Sp^k_{cp}}. \quad (15)$$

В дальнейшем Участники ранжируются по значению S^k и к Финальному этапу допускаются 4 (четыре) Команды, с наибольшими значениями S^k .

Приложение 1в

Алгоритм определения победителей Финального этапа

Преодоление Командой технологического барьера в Конкурсе фиксируется в случае, если СППВР данной Команды по результатам Испытаний Финального этапа демонстрирует показатели Se и Sp по каждой нозологии Конкурса, превышающие пороговые значения, указанные в п.5.5 Технического регламента Конкурса. Т.е. условия (9) истинны для каждой нозологии Конкурса.

Ранжирование Команд в лидерборде Финального этапа проводится по убыванию значения Интегрального критерия - комплексной метрики, представляемой несколькими последовательно проверяемыми условиями.

Пусть в финал конкурса вышли $2 \leq K_{\Phi} \leq 4$ Команды. По результатам Испытаний СППВР 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 в соответствии с порядком действий на стадии 1. По результатам испытаний на стадии 1 формируем сводную матрицу ошибок, в которой к результатам точности СППВР на Полуфинальном этапе добавляем результаты стадии 1:

$$TP^{\Phi} = TP^{пф} + TP; \quad TN^{\Phi} = TN^{пф} + TN; \quad FP^{\Phi} = FP^{пф} + FP; \quad FN^{\Phi} = FN^{пф} + FN. \quad (16)$$

Пересчитываем финальные значения точности Se и Sp СППВР участников по результатам Стадии 1. Для расчета нижних границ доверительных интервалов используется методика, представленная в приложении 1а. По результатам расчетов для Команды формируется таблица по каждой нозологии:

N п.п.	Метрика	Значение	Доверительный интервал		Технологический барьер
			CI_min ^Φ	CI_max ^Φ	
1	Se^{Φ} (TPR)	92,00%	84,78%	94,22%	82,00%
2	Sp^{Φ} (TNR)	87,7%	78,68%	96,12%	79,00%

Для Команд, чьи СППВР показывают значения точности, превышающие пороговые значения, указанные в п. 5.5 Технического регламента Конкурса (т.е. условия (9) истинны для каждой нозологии Конкурса) производится расчет обобщенного качества S^k :

Считается геометрические средние по чувствительности и специфичности, показанные по D-нозологиям:

$$Se_{cp}^k = \sqrt[6]{\prod_{d=1}^6 Se_{d}^k} \quad (17)$$

$$Sp_{cp}^k = \sqrt[6]{\prod_{d=1}^6 Sp_{d}^k}.$$

Производится расчет обобщенного качества S^k :

$$S^k = \sqrt[2]{Se_{cp}^k * Sp_{cp}^k}. \quad (18)$$

Значения S^k используются для дальнейшего расчета Интегрального критерия.

Далее, для показателя S рассчитывается дисперсия Var_S :

$$Var_S = \sqrt{\frac{S \cdot (1-S) + \frac{z_{1-\alpha}^2}{4 \cdot n_0}}{n_0}}. \quad (19)$$

Сравнение показателей S производится с помощью t-критерия Стьюдента. Для целей Конкурса полагаем разницу между значениями S^{j_1} и S^{j_2} (значения обобщенного качества для систем участников j_1 и j_2 соответственно) несущественной (значения совпадают), если значению t соответствует уровень $p < 0.05$ (рассчитанное значение t превышает табличное для $p=0.05$). Расчет значения t :

$$t = \frac{(S^{j_1} - S^{j_2}) \cdot \sqrt{2|N|}}{\sqrt{Var_{S^{j_1}} + Var_{S^{j_2}}}}. \quad (20)$$

Для расчета p используется количество степеней свободы, равное $2|N| - 2$, где $|N|$ - количество объектов в выборке N .

В качестве обоснованности в Интегральном критерии принимается статистическая обоснованность, а именно ширина разлета от нижнего значения доверительного интервала до верхнего предельного значения, равного единице. Чем меньше разлет, тем более обоснованным является результат СППВР. Для сравнения обоснованностей результатов СППВР используются нижние значения доверительных интервалов. Чем больше нижнее значение, тем более узким является разлет и, как следствие, большей обоснованность. В качестве примера, из двух систем участников большую обоснованность по Se демонстрирует та, чей $CI_{min}Se$ больше.

Интегральный критерий для сравнения решений двух участников рассчитывается последовательно, с обновлением лидерборда на каждом шаге. При этом, каждое

последующее сравнение производится только в случае равенства на предыдущем шаге (с округлением до второго знака после запятой). Расчет производится в следующем порядке:

1. Сравниваются значения обобщенного качества S^* .
2. Сравниваются значения обоснованности для Se .
3. Сравниваются значения обоснованности для Sp .
4. Сравниваются ресурсоемкости (включающие количество итераций).
5. Сравниваются Se осложнений.
6. Сравниваются Se сопутствующих заболеваний.
7. Сравниваются усредненное время постановки диагноза (на стадии 1 финала).

В случае, если ни одной СППВР не продемонстрировано преодоление Технологического барьера - производится проверка на Частичное преодоление Технологического барьера.

Пересчитываем финальные значения точности Se и Sp СППВР участников по результатам стадии 1. Для расчета нижних границ доверительных интервалов используется методика, представленная в приложении 1а. По результатам расчетов для Команды формируется таблица по каждой нозологии:

N п.п.	Метрика	Значение	Доверительный интервал		Технологический барьер
			CI_min ^Ф	CI_max ^Ф	
1	Se^{Φ} (TPR)	92,00%	84,78%	94,22%	82,00%
2	Sp^{Φ} (TNR)	87,7%	78,68%	96,12%	79,00%

Частичное преодоление технологического барьера определяется по показателю обобщенного качества S^* , рассчитанного отдельно для каждой нозологии (формулы 17,18). Таким образом, для каждой СППВР формируется массив из шести значений обобщенного качества (по каждой нозологии). Далее производится расчет Интегрального критерия для сравнения результатов СППВР отдельно по каждой нозологии Конкурса. Расчет Интегрального критерия производится по правилам, соответствующим полному преодолению технологического барьера. При этом, каждое последующее сравнение производится только в случае равенства на предыдущем шаге результатов нескольких СППВР (с округлением до второго знака после запятой). Расчет производится в следующем порядке:

1. Сравниваются значения обобщенного качества S^* .
2. Сравниваются значения обоснованности для Se .
3. Сравниваются значения обоснованности для Sp .
4. Сравниваются ресурсоемкости (включающие количество итераций).
5. Сравниваются Se осложнений.
6. Сравниваются Se сопутствующих заболеваний.
7. Сравниваются усредненное время постановки диагноза (на стадии 1 финала).

Ранжирование Команд по результатам Частичного преодоления технологического барьера производится в соответствии с нижеизложенной методикой:

Методика определения Частичного преодоления технологического барьера и ранжирования Команд:

Призовые места за Частичное преодоление могут быть присуждены 1-3 Командам в том случае, если не было зафиксировано полное преодоление Технологического барьера (ТБ).

В основу Частичного преодоления положены следующие утверждения:

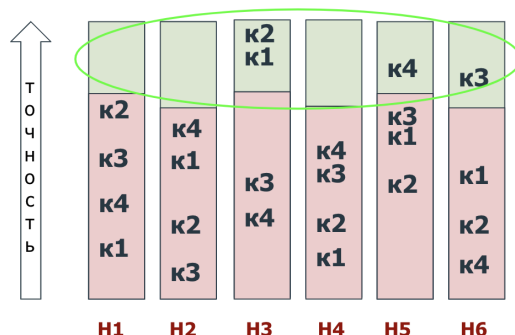
1. Команда должна продемонстрировать преодоление пороговых значения Se и Sp , соответствующих Технологическому барьеру, минимум по одной нозологии. Нозологии в рамках Конкурса считаются равнозначными.
2. В рамках каждой нозологии важна только наиболее точная СППВР, преодолевающая пороговые значения Se и Sp , соответствующих Технологическому барьеру.
3. Точность СППВР важнее количества нозологий, по которым данной СППВР продемонстрировано преодоление пороговых значения Se и Sp , соответствующих Технологическому барьеру. Количество нозологий, по которым данной СППВР продемонстрировано преодоление пороговых значения Se и Sp , соответствующих Технологическому барьеру не влияет на процесс ранжирования Команд в лидерборде Частичного преодоления.
4. В случае пороговых значения Se и Sp , соответствующих Технологическому барьеру по одной из нозологий, СППВР участвует в ранжировании по описанным правилам, независимо от точности по остальным нозологиям (вплоть до нулевых значений).

Фиксация Частичного преодоления Технологического барьера и ранжирование Команд осуществляется следующим образом:

Входные данные:

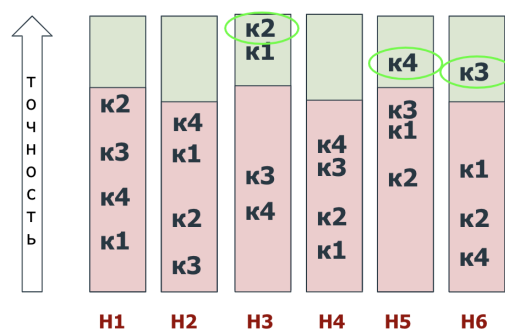
Команды K1, K2, K3, K4. Нозологии H1-H6

1. Определяются Команды, которые продемонстрировали превышение пороговых значений Se и Sp , соответствующих Технологическому барьеру минимум по одной нозологии. Для данных Команд в дальнейших расчетах используется метрика обобщенного качества S^* отдельно по каждой нозологии (на схемах представлена как “точность”).



В данном примере тремя наиболее точными признаны СППВР Команд K2, K1, K4. Но в соответствии с методикой Частичного преодоления, выбираются наиболее точные команды в разных нозологиях.

2. Определяем наиболее точные Команды в нозологиях. При этом используем абсолютные значения обобщенного качества S^* в каждой нозологии. Различия между пороговыми значениями Технологических барьеров в разных нозологиях не учитываются.



В данном примере тремя претендентами на Частичное преодоление признаны СППВР Команд К2, К4, К3.

3. Приз за частичное преодоление присуждается следующим образом: I место - К2; II место - К4; III место - К3 (по уменьшению значения обобщенного качества S^k).

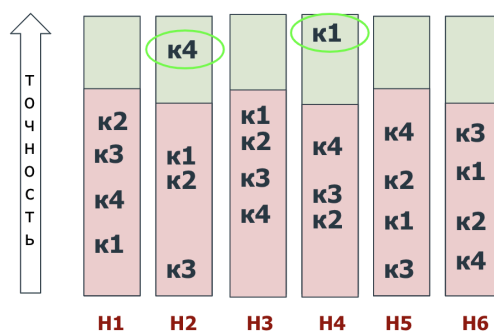
Возможные ситуации, в которых происходит изменение вышеизложенной методики:

1. Зарегистрировано преодоление ТБ менее чем в трех нозологиях.



Решение: Определяются наиболее точные СППВР в разных нозологиях. Далее, определяются наиболее точные без учета нозологий. Приз за Частичное преодоление присуждается следующим образом: I место - К1; II место - К4; III место - К2.

2. Зарегистрировано преодоление ТБ менее чем в трех нозологиях, при этом ТБ преодолели менее трех команд.



ИЛИ

AI'm Doctor



Решение: Ранжирование по точности без учета нозологий. Два призовых места.

Приложение 2

Типы Сессий обмена с Платформой в рамках Отборочного этапа Конкурса

Тренировочная Сессия

Тип Сессии - training

Ответ должен приходить 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 приходить 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.

Приложение 3

Типы Сессий обмена с Платформой в рамках Полуфинального этапа Конкурса

Общая информация

Во время проведения сессий Полуфинального этапа конкурса AI'm Doctor, СППВР, в рамках одного Эпикриза, будет получать его поочередно в двух версиях:

- вначале - v3, который может содержать сокращенную номенклатуру произведенных исследований;
- затем, через 30 сек, - v2, с полным набором произведенных медицинских исследований.

Для каждого выданного в обработку файла (v2 и v3), СППВР должна будет вернуть на Платформу ответ, который может включать в себя указания:

- основного заболевания заключительного диагноза в виде кода МКБ-10 (обязательно должен присутствовать в ответе, в единственном числе для каждого выданного на обработку файла. В противном случае, ответ будет автоматически считаться невалидным);
- осложнения основного заболевания в виде кода МКБ-10. Указание осложнения в ответе имеет значение только для ответа на версию v2 (в Эпикризе может отсутствовать или быть в количестве более одного);
- сопутствующих заболеваний в виде кода МКБ-10. Указание сопутствующего заболевания в ответе имеет значение только для ответа на версию v2 (в Эпикризе может отсутствовать или быть в количестве более одного).

Ответ СППВР по версии v3 Эпикриза влияет исключительно на ресурсоемкость.

Ответ СППВР по версии v2 Эпикриза влияет на значение обобщенного качества S^k .

Формат ответа

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

1. При указании Основного заболевания объект должен иметь вид:
 - decorCode - со значением "diagnosisMain"
 - code - со значением одного кода МКБ-10, либо "another" (если вы указываете что пациент не болен ни одной из 6ти нозологий конкурса), либо пустая строка ("") если Участник не считает возможным поставить диагноз.

В случае, если Участник не укажет значение ключа "diagnosisMain" и пришлет в ответе объект вида:

```
"decorCode": "diagnosisMain",
"code": ""
```

Платформа засчитает такое указание Основного заболевания валидным, однако, ответ будет засчитан неверным

Ограничение

В массиве, который будет прислан как ответ по конкретному Эпикризу, должен содержаться только один объект с decorCode “diagnosisMain”. В случае отсутствия такового объекта в массиве, либо если их будет больше одного для одного решения - ответ будет считаться невалидным.

2. При указании осложнений Основного заболевания (имеет значение только при ответе на v2), каждое осложнение должно быть оформлено отдельным объектом, входящим в корневой массив ответа. Объект вида:
 - decorCode - со значением “attendDisease”
 - code - со значением одного кода МКБ-10

В массиве может быть несколько объектов, указывающих на Осложнения Основного заболевания.

Количество объектов массива со значением “attendDisease” в ключе decorCode должно быть не более десяти, иначе, ответ будет считаться невалидным.

3. При указании сопутствующих заболеваний, каждое заболевание должно быть оформлено отдельным объектом, входящим в массив. Объект вида:
 - decorCode - со значением “diagnosisSup”
 - code - со значением одного кода МКБ-10

В массиве может быть несколько объектов, указывающих на Сопутствующие заболевания.

Количество объектов массива со значением “diagnosisSup” в ключе decorCode должно быть не более десяти, иначе, ответ будет считаться невалидным.

Типы сессий:

Тренировочная сессия

Расчет по данному виду сессии не производится, после проведения сессии, в ЛК Участника публикуются логи сессии по обмену данными с Платформой.

Параметризация сессии:

Параметр	Комментарий
count	Настройка количества эпикризов - Участник может выставить значение данного параметра в промежутке от 1 до 2000 эпикризов

	Ознакомительном датасете случайным образом, вне зависимости нозологии. По умолчанию стоит значение 100 эпикризов. Параметр count характеризует количество эпикризов, которые будут выданы в рамках сессии. Таким образом, количество файлов за сессию будет выдано в два раза больше, чем указано в параметре count. Файлы будут выдаваться поочередно v3, затем v2 версии одного эпикриза
timeout	Настройка интервала времени выдачи файлов - Участник может выставить значение данного параметра в промежутке от 10 до 30 секунд. По умолчанию стоит значение 30 сек. Параметр характеризует промежуток времени, между выдачей файлов с одной стороны Платформы Участнику. Т.е. время между выдачей v3 и v2, затем следующего v3 и т.д

Расчетная тренировочная сессия

Расчет результатов по данному виду сессий будет происходить на Платформе два раза в сутки, затем публиковаться в ЛК Участников Конкурса.

Ограничение на проведение сессии данного вида - не больше **10** раз в сутки для одного Участника.

Параметризация сессии:

Параметр	Комментарий
count	Для данного типа сессии - отсутствует. Используется фиксированное значение 70 эпикризов на нозологию
timeout	Настройка интервала времени выдачи файлов - Участник может выставить значение данного параметра в промежутке от 10 до 30 секунд. По умолчанию стоит значение 30 сек. Параметр характеризует промежуток времени, между выдачей файлов с одной стороны Платформы Участнику. Т.е. время между выдачей v3 и v2, затем следующего v3 и т.д
Выбор нозологии	Настройка выдачи эпикризов на проверку соответствия конкретной нозологии. По умолчанию раздаются все нозологии Конкурса

Испытательная сессия

Расчет результатов по данному виду сессии будет происходить на Платформе, после проведения Испытаний. Публикация результатов будет производиться в обозначенное Регламентом Конкурса время в публичном Лидерборде.

Примеры ответов

Пример 1

Валидный ответ с указанием Основного заболевания:

```
[
  {
    "decorCode": "diagnosisMain",
    "code": "C34"
  }
]
```

Пример 2

Валидный ответ с пропуском Основного заболевания:

```
[
  {
    "decorCode": "diagnosisMain",
    "code": ""
  }
]
```

Пример 3

Валидный ответ с указанием Основного заболевания, сопутствующего заболевания и Осложнения:

```
[
  {
    "decorCode": "diagnosisMain",
    "code": "another"
  },
  {
    "decorCode": "diagnosisSup",
    "code": "C34"
  },
  {
    "decorCode": "attendDisease",
    "code": "C34"
  }
]
```

Пример 4

Не валидный ответ без наличия Основного заболевания:

```
[
  {
    "decorCode": "diagnosisSup",
    "code": "C34"
  },
  {
    "decorCode": "attendDisease",
    "code": "C34"
  }
]
```

Пример 5

Не валидный ответ с указанием более 10 значений Осложнений:

```
[
  {
    "decorCode": "diagnosisMain",
    "code": "C34"
  },
  {
    "decorCode": "attendDisease",
    "code": "C34"
  },
  {
    "decorCode": "attendDisease",
    "code": "C34"
  },
  {
    "decorCode": "attendDisease",
    "code": "C34"
  },
  {
    "decorCode": "attendDisease",
    "code": "C34"
  },
  {
    "decorCode": "attendDisease",
    "code": "C34"
  }
]
```

```
    "code": "C34"
  },
  {
    "decorCode": "attendDisease",
    "code": "C34"
  },
  {
    "decorCode": "attendDisease",
    "code": "C34"
  },
  {
    "decorCode": "attendDisease",
    "code": "C34"
  },
  {
    "decorCode": "attendDisease",
    "code": "C34"
  }
]
```

Приложение 4

Таблица исследований Финального этапа

Инструментальные исследования

№	Код запроса исследования	Исследование или группа исследований в рамках Конкурса	Код медуслуги, возвращаемой в ответ на запрос
1	A03.09	Бронхоскопия	A03.09.001
2	A03.16	Эзофагогастродуоденоскопия	A03.16.001
3	A03.30	Эндоскопическое исследование внутренних органов	A03.30.006
4	A04.09	Ультразвуковое исследование плевральной полости	A04.09.001
5	A04.10	Эхокардиография	A04.10.002
6	A04.12	УЗИ вен нижних конечностей	A04.12.006.002
7	A04.16	УЗИ органов брюшной полости	A04.16.001
8	A04.21	УЗИ предстательной железы трансабдоминальное	A04.21.001
9	A04.28	УЗИ почек и надпочечников трансабдоминальное	A04.28.001
10	A05.10	Регистрация электрокардиограммы	A05.10.006
11	A05.10	Холтер	A05.10.008
12	A05.23	MPT головного мозга	A05.23.009.001
13	A06.03	Рентгенография грудины	A06.03.024
14	A06.08	Рентгенография/КТ околоносовых пазух	A06.08.003
			A06.08.007.003
15	A06.09	Рентген/КТ/флюорография легких	A06.09.001
			A06.09.005
			A06.09.006.001
			A06.09.007
			A06.09.007.002
			A06.09.008.001
16	A06.12	Антиография	A06.12.001.001
			A06.12.049
			A06.12.010
17	A06.16	Рентген пищевода	A06.16.001.002
18	A06.20	Мамография	A06.20.004.006
19	A06.30	КТ живота	A06.30.005
20	A07.09	Сцинтиграфия лёгких перфузионная	A07.09.003
21	A12.09	Спирометрия и спирография	A12.09.001
			A12.09.002
			A12.09.002.001
			A12.09.002.002
22	A12.09.005	Пульсоксиметрия	A12.09.005

Лабораторные исследования

№	Код запроса исследования	Исследование в рамках Конкурса	Код медуслуги, возвращаемой в ответ на запрос
1	B03.005.006	Коагулограмму с определением протромбинового времени, международного нормализованного отношения и активированного частичного тромбопластинового времени	B03.005.006
2	B03.016.003	Развернутый клинический анализ крови	B03.016.003
3	B03.016.004	Биохимический анализ крови	B03.016.004
4	B03.016.006	Анализ мочи	B03.016.006
5	B03.016.011	Исследование газов артериальной крови	B03.016.011
6	A09.05.209	Прокальцитонин	A09.05.209
7	A11.01.003	Диаскинтест	A11.01.003.001
8	A26.09.010	Микроскопическое исследование мокроты на наличие бактерий	A26.09.010
9	A26.30.028	GeneXpert БАЛ	A26.30.028