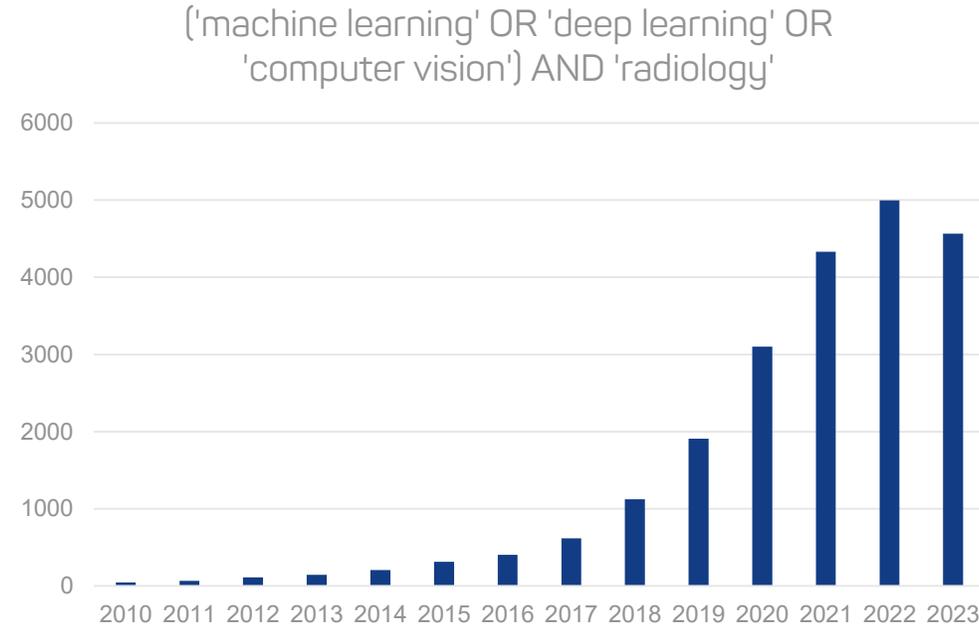
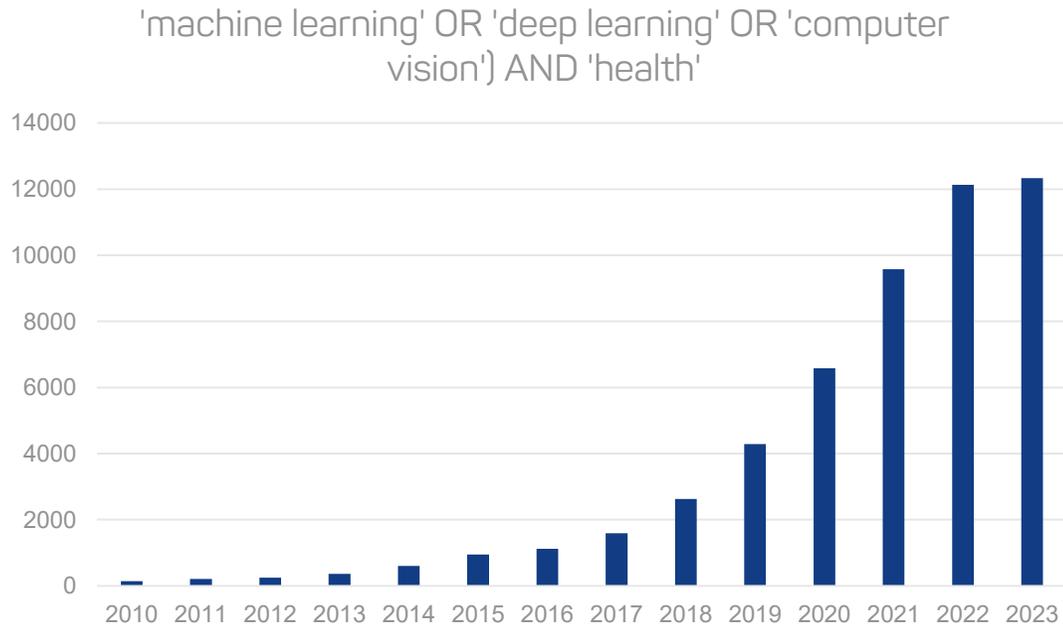


Оценка качества ПО на основе технологий искусственного интеллекта в лучевой диагностике на всех этапах жизненного цикла

АРЗАМАСОВ Кирилл Михайлович,
к.м.н., руководитель отдела медицинской информатики, радиомики и радиогеномики

Москва, 2024 г.

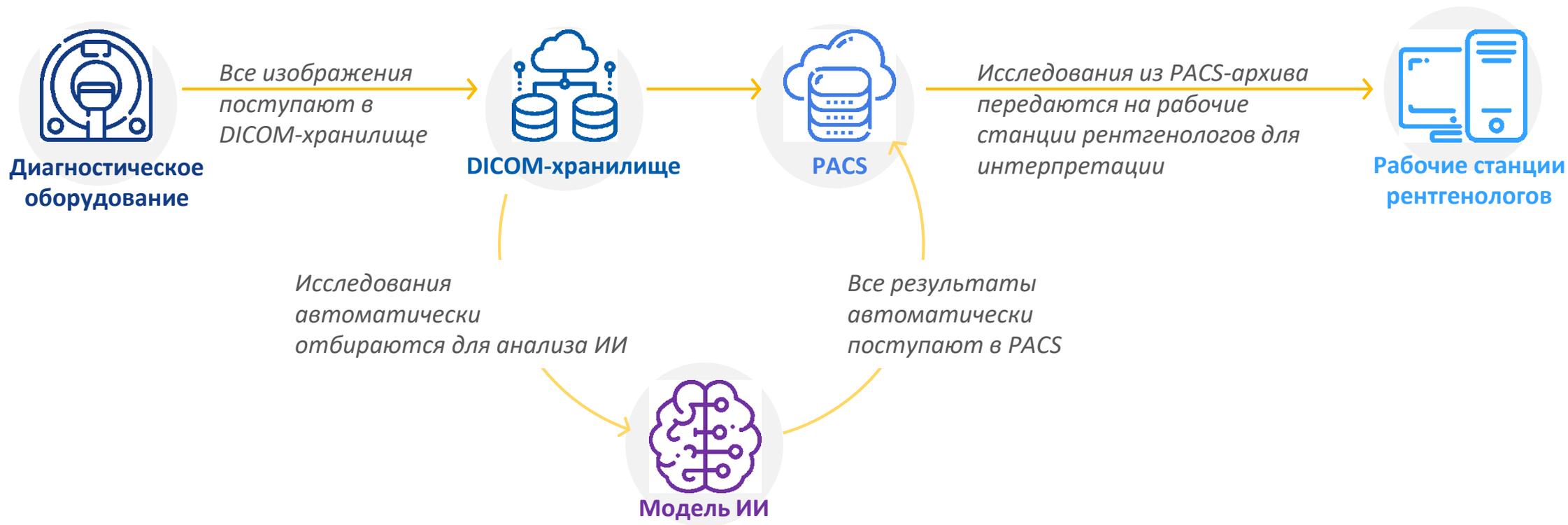


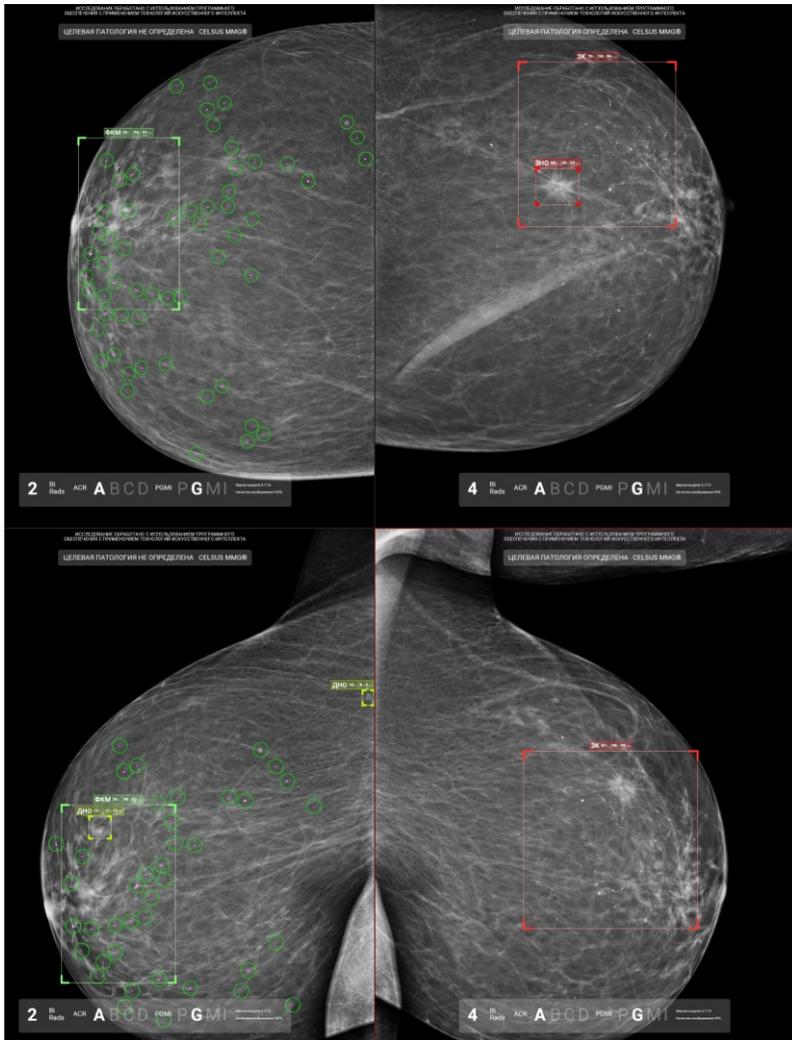


Новое поколение **средств автоматизации** производственных процессов в сфере здравоохранения - **технологии искусственного интеллекта** (ТИИ). Количество зарегистрированных (FDA) ИИ-сервисов в лучевой диагностике ~100, в РФ (медицинское изделие) > 22.

Национальной стратегией развития искусственного интеллекта в России до 2030 г. предполагается **использование ТИИ в сфере здравоохранения** - профилактические обследования и диагностика.

РАБОЧИЙ ПРОЦЕСС ПРИМЕНЕНИЯ ИИ





Пример реализации DICOM SC

ОПИСАНИЕ:

Сторона L: Структура молочной железы по ACR - A. Молочная железа практически полностью жировой плотности. Маммография обладает высокой чувствительностью.

Проекция: CC Определены:

Злокачественный кальцинат, вероятность находки - 0,70, вероятность злокачественности - 0,70, площадь 7740 кв. мм, высота = 90 мм, ширина = 86 мм. Злокачественное новообразование, вероятность находки - 0,68, вероятность злокачественности - 0,68, площадь 460 кв. мм, высота = 20 мм, ширина = 23 мм.

Проекция: MLO Определены:

Злокачественный кальцинат, вероятность находки - 0,67, вероятность злокачественности - 0,67, площадь 9120 кв. мм, высота = 96 мм, ширина = 95 мм.

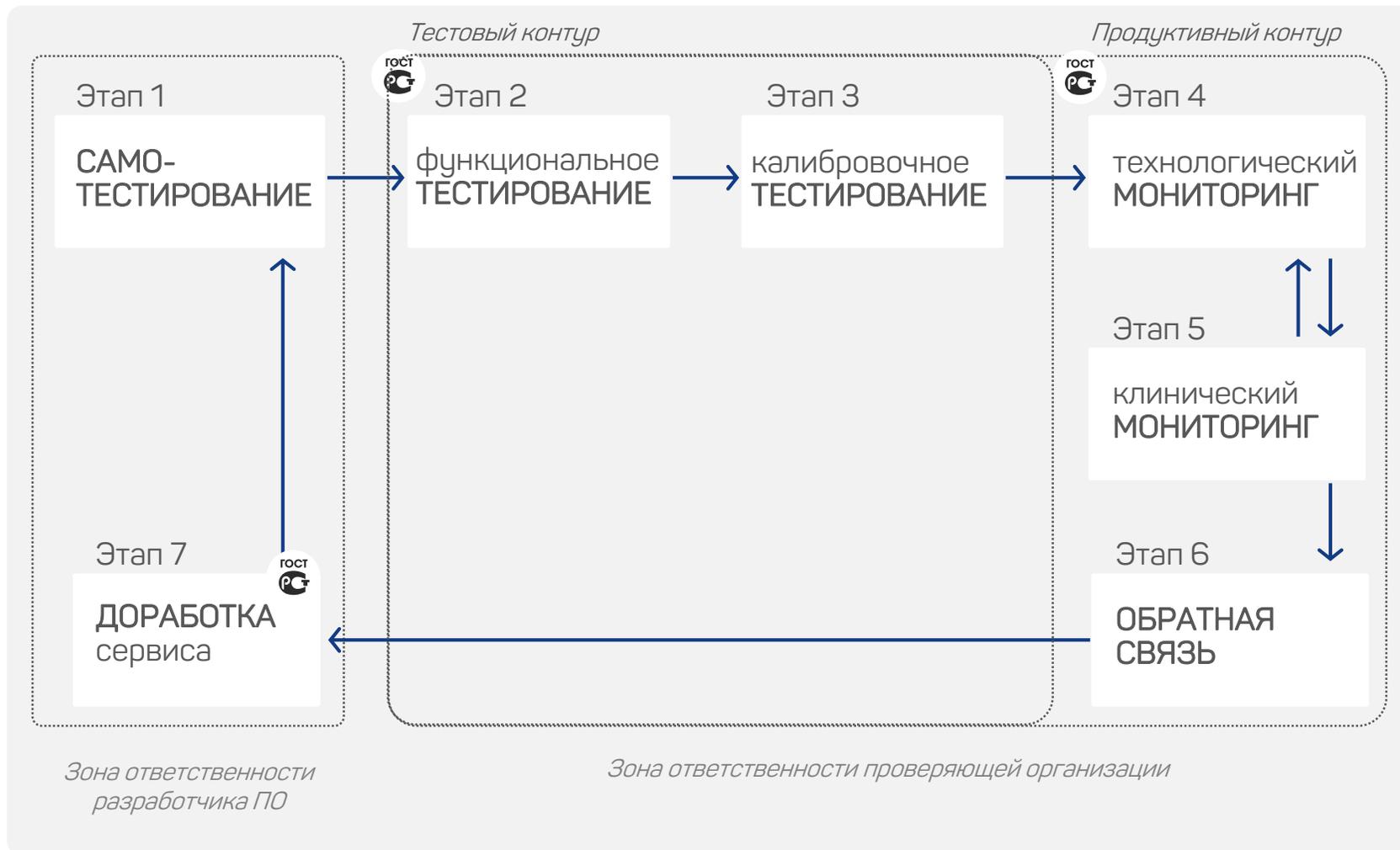
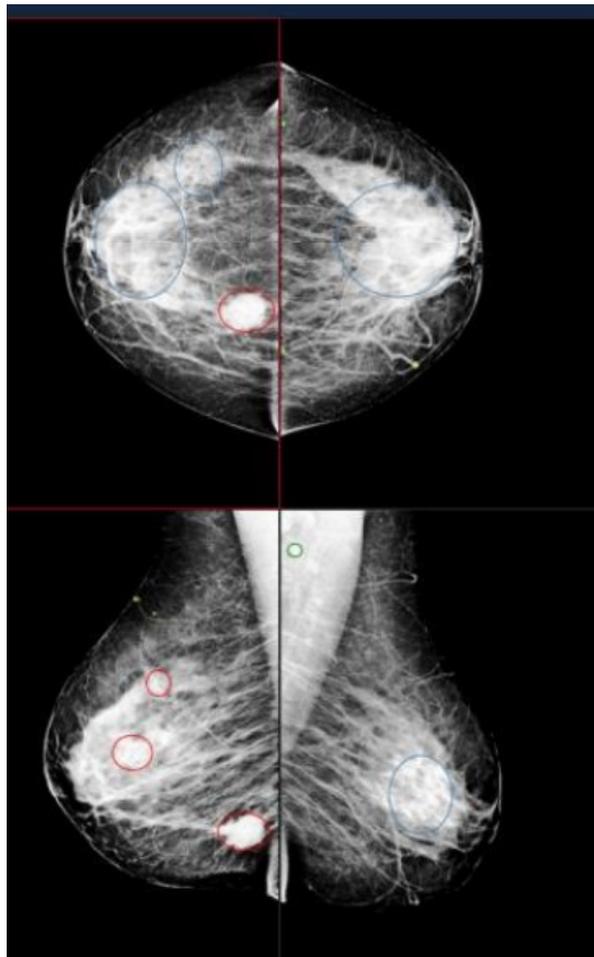
ЗАКЛЮЧЕНИЕ:

Оценка категории по диагностической шкале BI-RADS: для правой молочной железы BI-RADS 2, для левой молочной железы BI-RADS 4

Скрининг оценка категории: для правой молочной железы BI-RADS 2, для левой молочной железы BI-RADS 0

Пример DICOM SR

МЕТОДОЛОГИЯ ТЕСТИРОВАНИЯ И МОНИТОРИНГА ПО



Оценка осуществляется на основании сопоставления результата работы ИИ-сервиса и заключения врача.

Заключение врача по бинарной шкале:

0 – нет целевой патологии,

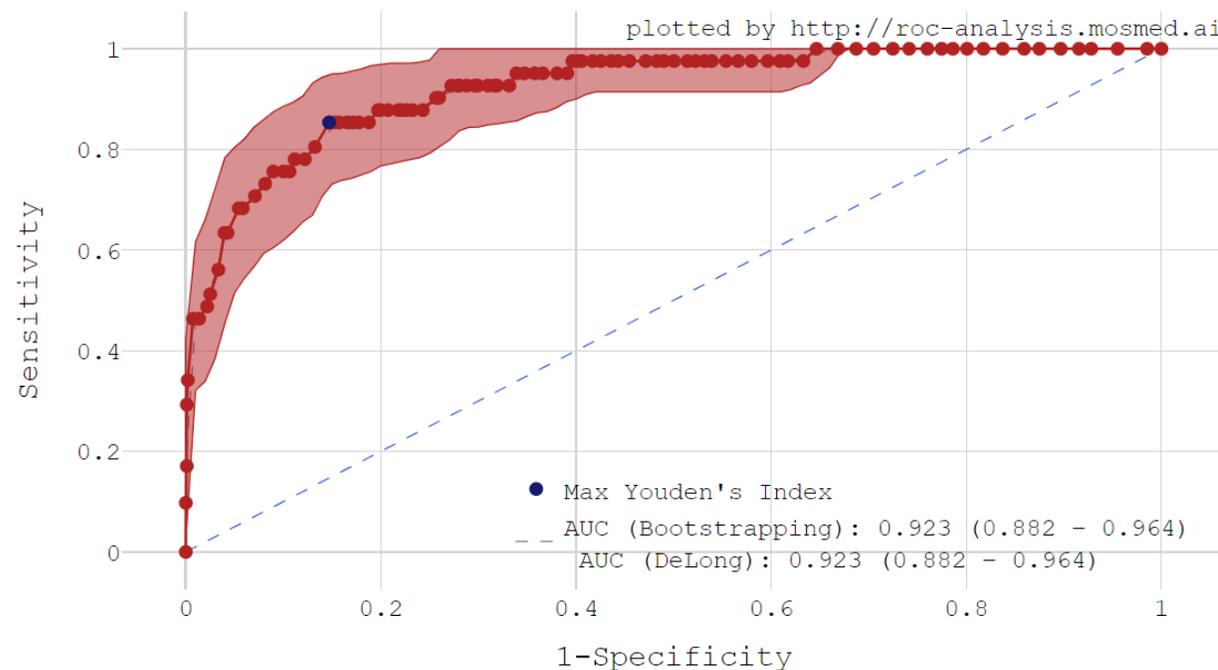
1 – есть целевая патология.

Результат работы ИИ-сервиса – вероятность наличия патологии.

Параметр, на основании которого осуществляется оценка ИИ-сервиса – площадь под характеристической кривой (AUC).

Оптимальным считается AUC **>0,81**

[МР «Клинические испытания программного обеспечения на основе интеллектуальных технологий (Лучевая диагностика)»]

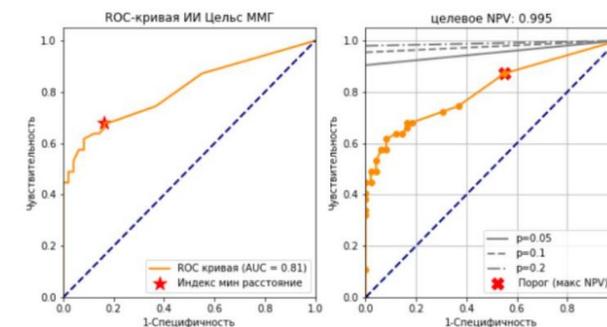


Характеристическая кривая ИИ-сервиса

КАЛИБРОВОЧНОЕ ТЕСТИРОВАНИЕ

Показатели диагностической точности ПО, определенные в результате калибровочного тестирования

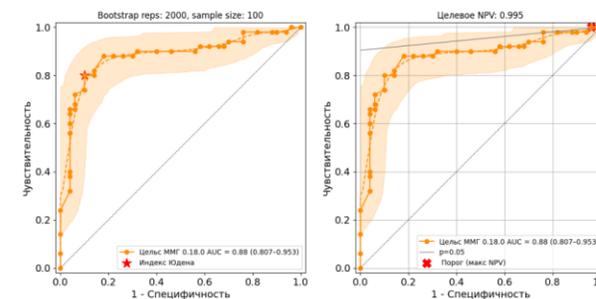
	КТ 1 (2020 г.)	КТ 2 (2021 г.)	КТ 3 (2022 г.)	Результат
Площадь под ROC-кривой (AUC)	0,81	0,94	0,91	+12,35%
Точность	0,76	0,85	0,89	+17,11%
Чувствительность	0,68	0,92	0,85	+25%
Специфичность	0,84	0,78	0,93	+10,71%



Оптимальный порог
(минимальное расстояние): 0.6

Оптимальный порог (макс NPV): 0.56
пре-тестовая вероятность: 0.05

ROC-кривые по данным 1-го КТ



Оптимальный порог (индекс Юдена): 77.0

Оптимальный порог (макс NPV): 3.0
пре-тестовая вероятность: (0.05,)

ROC-кривые по данным 3-го КТ

ТЕХНОЛОГИЧЕСКИЕ ДЕФЕКТЫ

Технологический мониторинг – оценка соответствия результатов работы ИИ-сервисов базовым функциональным требованиям, а также нормативно-правовым актам по Эксперименту.

Основные параметры, подлежащие мониторингу:

1. **Наличие результатов** работы ИИ-сервиса (текстовое описание и дополнительное изображение, содержащее маркировку патологических областей)
2. **Время обработки** исследования (<6,5 мин.)
3. **Корректно выбранное** для обработки исходное изображение и др.

Предельно допустимый процент исследований с технологическими дефектами регламентирован приказом ДЗМ и соответствует 10%. Технически стабильный ИИ-сервис имеет минимальное количество исследований с технологическими дефектами.

Перечень технологических дефектов:

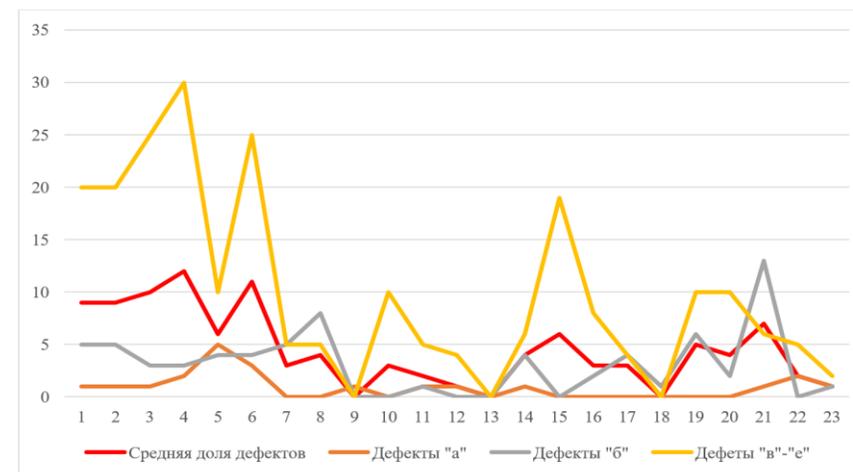
- в1 – отсутствие дополнительной серии
- в2 – отсутствие DICOM SR
- в3 – наличие 2-х и более DICOM SR
- в4 – отсутствие названия Сервиса
- в5 – отсутствие сведений о версии Сервиса

- г1 – изображения в дополнительной серии обрезаны
- г2 – яркость/контрастность дополнительной серии не соответствует оригинальному изображению
- г3 – проанализированы не все необходимые изображения
- г4 – отсутствие предупреждающей надписи: «Только для использования в исследовательских/научных целях»
- г5 – изменение оригинальной серии исследования

- д1 – разметка за пределами целевого органа
- д2 – проанализирована некорректная анатомическая область, проекция или серия

Показатели диагностической точности и доли дефектов ПО
по результатам двухлетнего мониторинга

	Начало мониторинга (2020 г.)	Конец мониторинга (2022 г.)	Изменение
Значение площади под ROC-кривой (AUC)	0,945 (0,849-1,048)	0,968 (0,926-1,008)	+2,43%
точность	0,9 (0,769-1,0)	0,950 (0,902-0,998)	+5,55%
чувствительность	0,857 (0,598-1,0)	1 (1,0-1,0)	+16,69%
специфичность	0,923 (0,778-1,0)	0,949 (0,900-0,998)	+2,82%
Дефекты «а»	1%	1%	0%
Дефекты «б»	5%	1%	-4%
Дефекты «в-е»	20%	2%	-18%



ОТЧЕТ ПО МОНИТОРИНГУ ТЕХНОЛОГИЧЕСКИХ ПАРАМЕТРОВ РАБОТЫ СЕРВИСА

- Компания-поставщик ИИ-сервиса: ООО "Медицинские скрининг системы"
Наименование ИИ-сервиса: Цельс
Компания-производитель ИИ-сервиса: ООО "Медицинские скрининг системы"
Идентификатор ИИ-сервиса в эксперименте: ЦЕСММГ
- Отчетный период: 05.04.2021 — 14.04.2021
- Вид исследований: ММГ
- Клиническая цель: рак молочной железы
- Общее количество исследований:
5.1. Направленных на анализ сервису за отчетный период согласно выгрузке из ЕРИС ЕМИАС, шт. 4022
5.2. Из них уникальных согласно выгрузке из ЕРИС ЕМИАС, шт. * 4022
- Количество исследований, прошедших контроль, шт. 4022
6.1. Прошедших ручной контроль, шт. 20
- Количество исследований, содержащих дефекты:
7.1. Содержащих технологический дефект «а», шт., приложение 1 35
7.2. Содержащих технологический дефект «б», шт., приложение 2 124
7.3. Содержащих технологические дефекты «в-е», шт., приложение 3 5
- Удельный вес исследований:
8.1. Содержащих технологический дефект «а» относительно 3898 исследований, % 1
8.2. Содержащих технологический дефект «б» относительно 4022 исследований, % 3
8.3. Содержащих технологические дефекты «в-е» относительно 20 исследований, % 25
- Количество исследований без дефектов, шт. 3858
- Решение:
РАБОТА СЕРВИСА ПРИОСТАНОВЛЕНА 14.04.2021
ПО РЕЗУЛЬТАТАМ МОНИТОРИНГА ЗА ПРЕДШЕСТВУЮЩИЙ ПЕРИОД
- Примечания: Уд. вес тех. деф. «в-е» > 10 %.

Дата оформления отчета: 06.07.2022
ФИО ответственного лица: отчет сформирован автоматически



* автоматизированная выборка из 80 исследований, отобранных случайным образом с обогащением в сторону исследований с патологией (до 60 исследований из выборки)

Объем выборки рассчитан при использовании подхода, основанного на проверке статистических гипотез

ОБРАТНАЯ СВЯЗЬ И ДОРАБОТКА

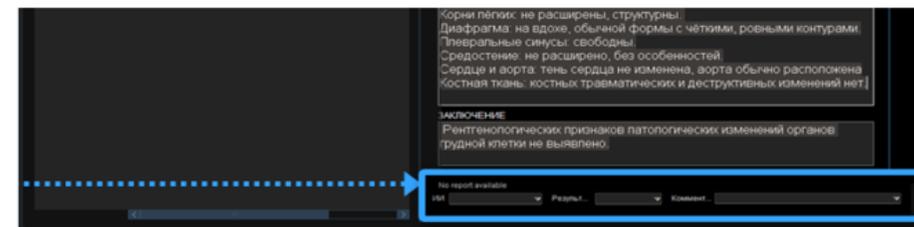
Всего было выполнено **18** доработок.

Были дополнены функции определения: PGM1, плотность МЖ по шкале ACR, шкала BI-RADS.

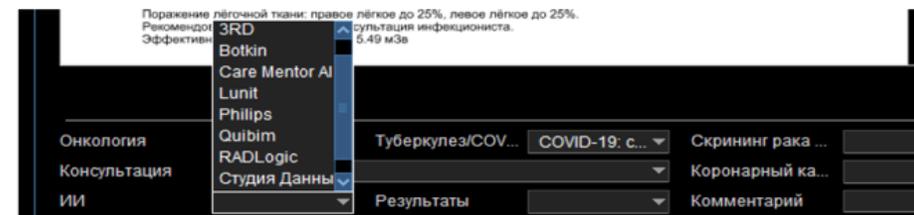
Изменение проспективных показателей диагностической точности

	2020 г.	2022 г.
AUC-ROC	0,81	0,91
Точность	0,76	0,89
Чувствительность	0,68	0,85
Специфичность	0,84	0,93
Доля дефектов «в-е»	20%	2%

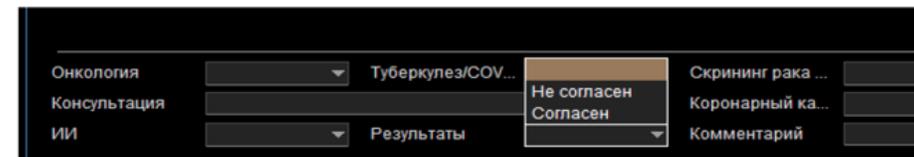
Поле обратной связи



Выберите сервис



Выберите «Согласен»



Форма обратной связи в радиологической информационной сети

При внедрении программного обеспечения (ПО) на основе технологий искусственного интеллекта (ТИИ) в практическое здравоохранение, важно сделать выбор в пользу зрелых решений.

Зрелым ПО на основе ТИИ является технически стабильное ПО, имеющее высокие метрики диагностической точности.

Цель настоящей работы: разработать и апробировать методологию оценки зрелости программного обеспечения на основе ТИИ для сферы здравоохранения.

Тыров И.А., Васильев Ю.А., Арзамасов К.М.,
Владзимирский А.В., Шулькин И.М., Омелянская О.В.,
Четвериков С.Ф..

Оценка зрелости технологий искусственного интеллекта для здравоохранения: методология и ее применение на материалах московского эксперимента по компьютерному зрению в лучевой диагностике.

Врач и информационные технологии. 2022; 4: 76-92. doi:
10.25881/18110193_2022_4_76.



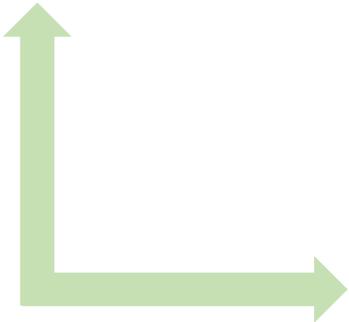
Оценка зрелости технологий искусственного интеллекта для здравоохранения: методические рекомендации / сост. Ю. А. Васильев, А. В. Владзимирский, О. В. Омелянская [и др.] // Серия «Лучшие практики лучевой и инструментальной диагностики». – Вып. 123. – М. : ГБУЗ «НПКЦ ДиТ ДЗМ», 2023. – 22 с.

КАЧЕСТВЕННАЯ СОСТАВЛЯЮЩАЯ МАТРИЦЫ

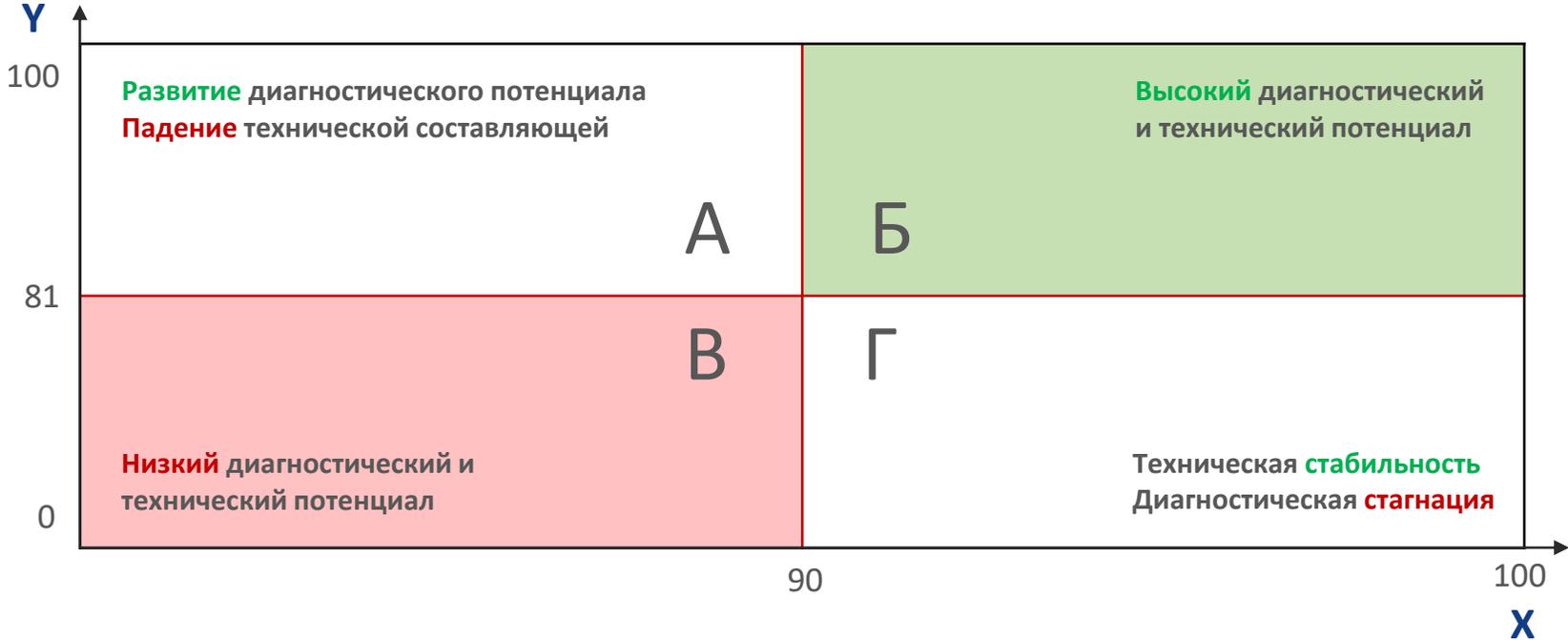
Ось OY – клиническая оценка
(проспективный ROC AUC * 100)

Граница «81» – горизонтальная линия с граничным значением для клинической значимости параметра «ROCAUC» в соответствии с методическими рекомендациями (№43), равное «0,81 * 100»

Повышение
диагностической
точности



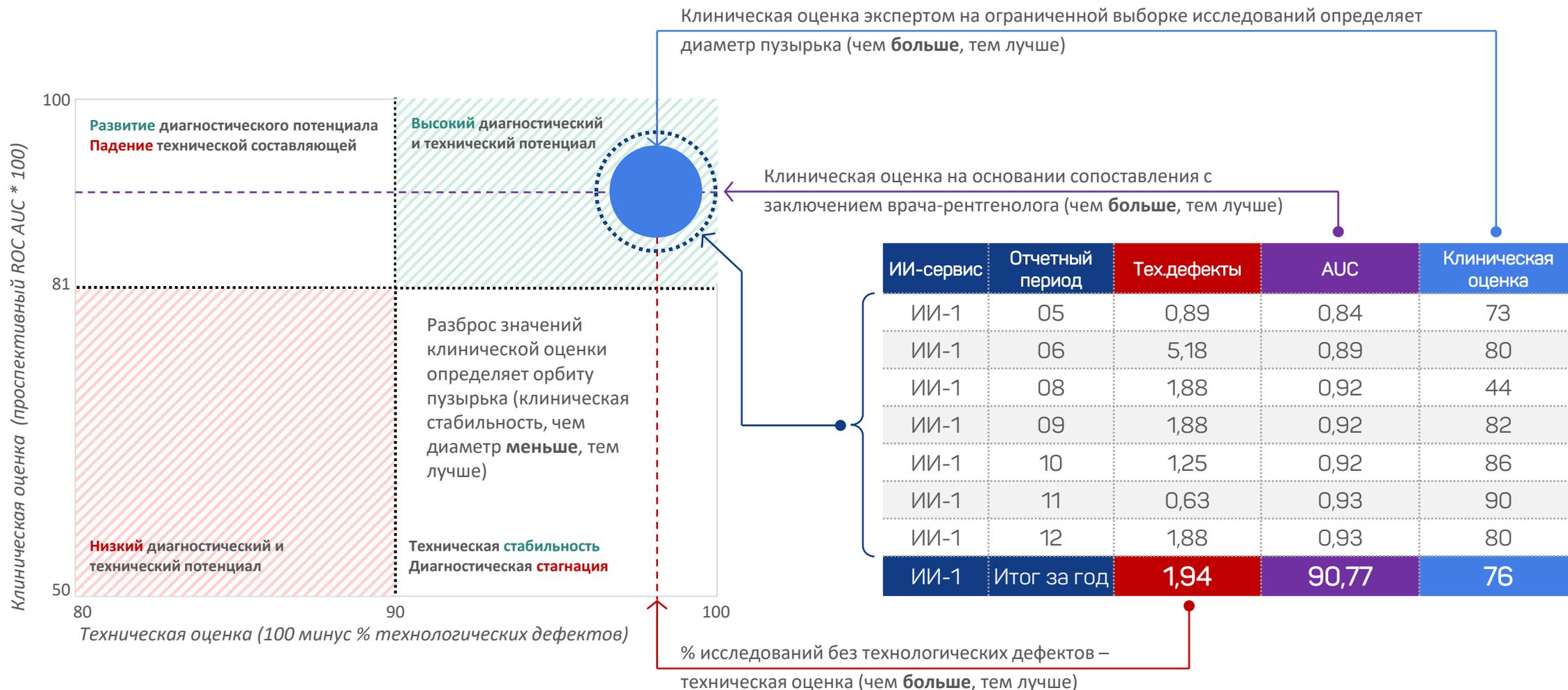
Повышение технической
стабильности



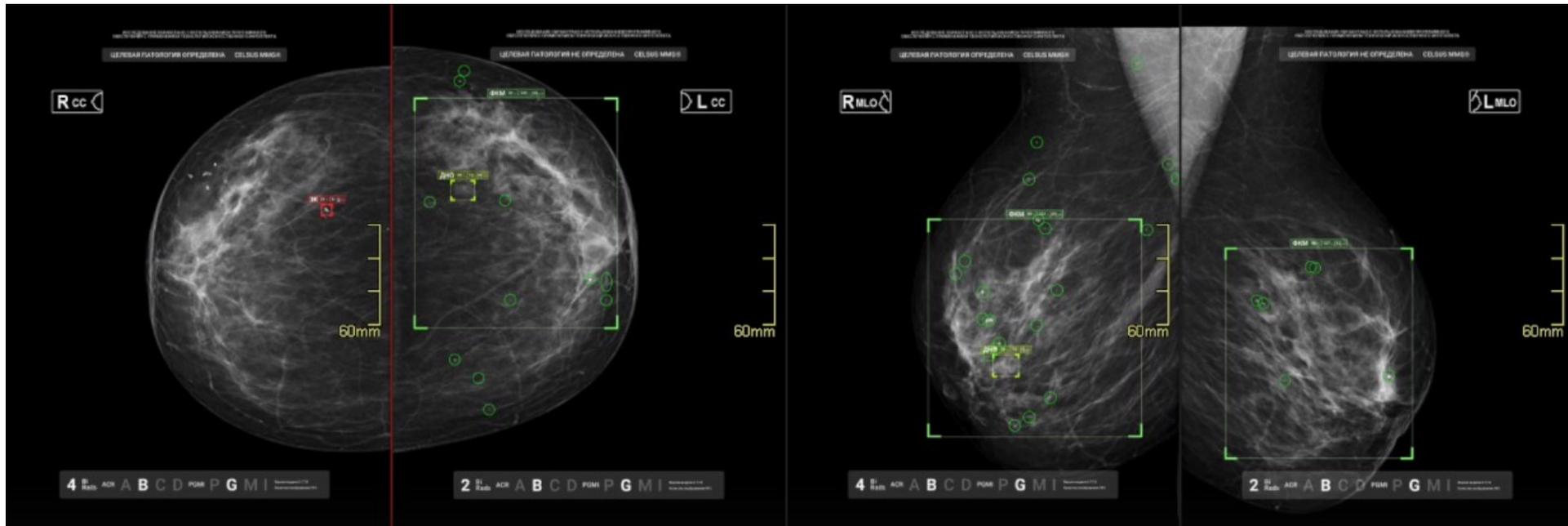
Ось OX – техническая оценка (100 минус % технологических дефектов)

Граница «90» – вертикальная линия отражает отметку 10% технологических дефектов, в соответствии с Приказом ДЗМ

МАТРИЦА ЗРЕЛОСТИ



Методология тестирования и мониторинга позволила повысить показатели диагностической точности ПО на основе ТИИ и как следствие - к повышению его практической значимости (первая медицинская услуга «Описание и интерпретация данных маммографического исследования с использованием искусственного интеллекта»).



Дополнительная серия изображений, созданная ПО, на котором размечены зоны фиброзно-кистозной мастопатии, доброкачественное и злокачественное образования

БЛАГОДАРИМ ЗА ВНИМАНИЕ



✉ npcmr@zdrav.mos.ru

☎ +7 (495) 276 - 04 - 36

🌐 telemedai.ru