



Методология матрицы зрелости ИИ-сервисов 2023 г.



Актуальность. Цель.



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.





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:

- v1 – отсутствие дополнительной серии
- v2 – отсутствие DICOM SR
- v3 – наличие 2-х и более DICOM SR
- v4 – отсутствие названия Сервиса
- v5 – 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



Клиническая оценка – сопоставление результата работы ИИ-сервиса и заключения врача.

Заключение врача по бинарной шкале:

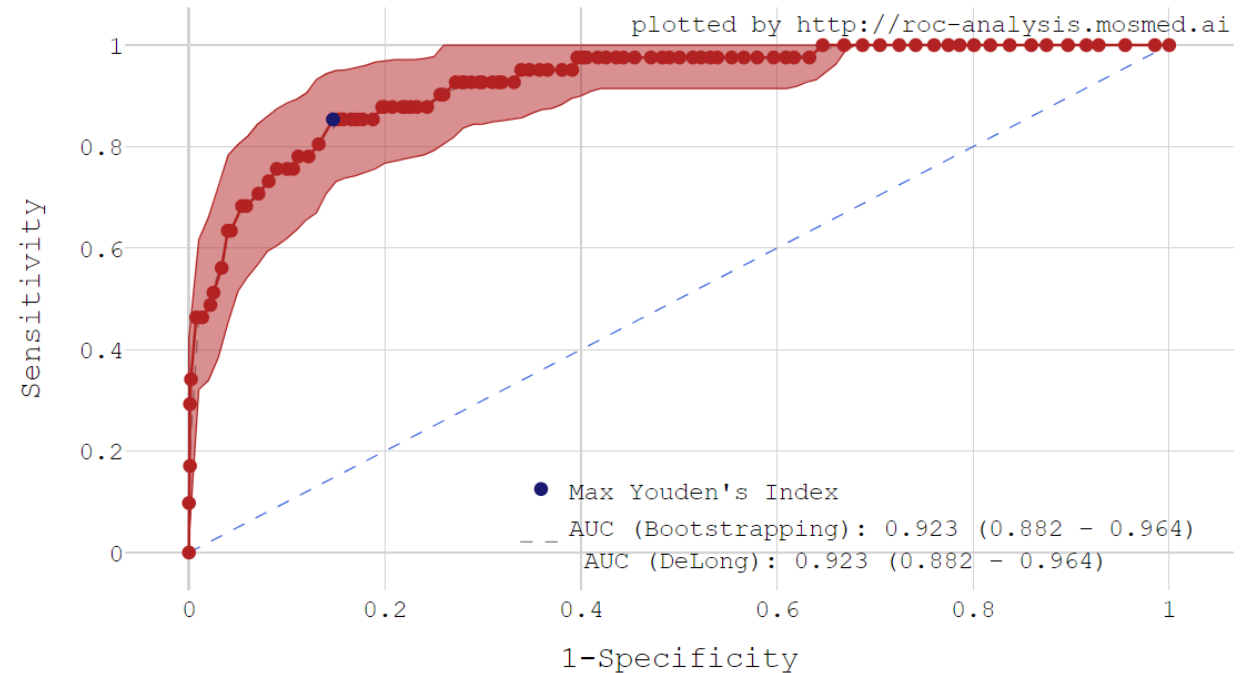
0 – нет целевой патологии,
1 – есть целевая патология.

Результат работы ИИ-сервиса – вероятность наличия патологии.

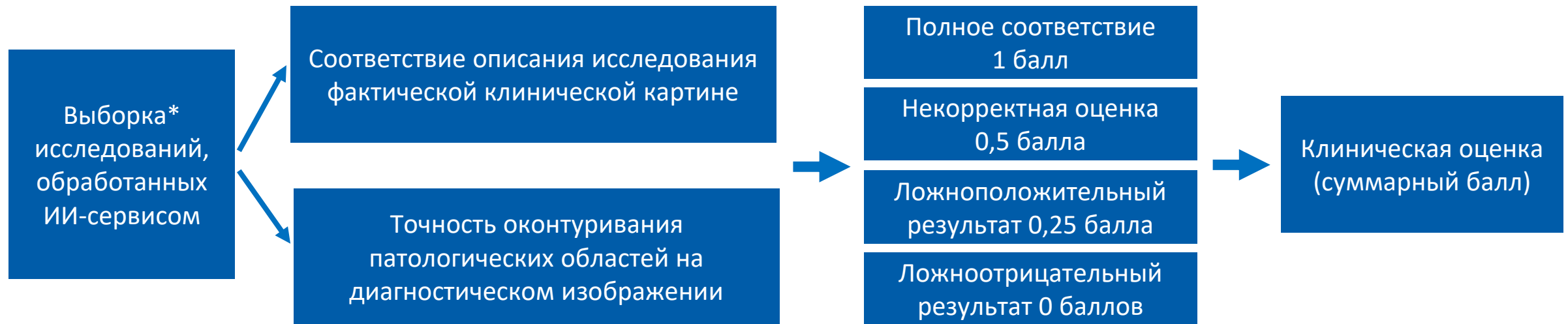
Параметр, на основании которого осуществляется оценка ИИ-сервиса – площадь под характеристической кривой (AUC).

Оптимальным считается AUC **>0,81**

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



Характеристическая кривая ИИ-сервиса



* 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 (научная статья в печати)



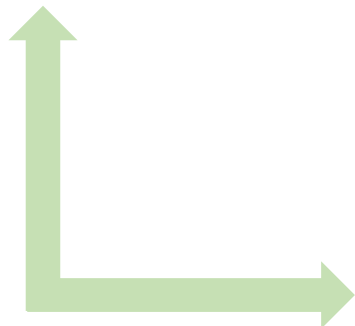
Качественная составляющая матрицы



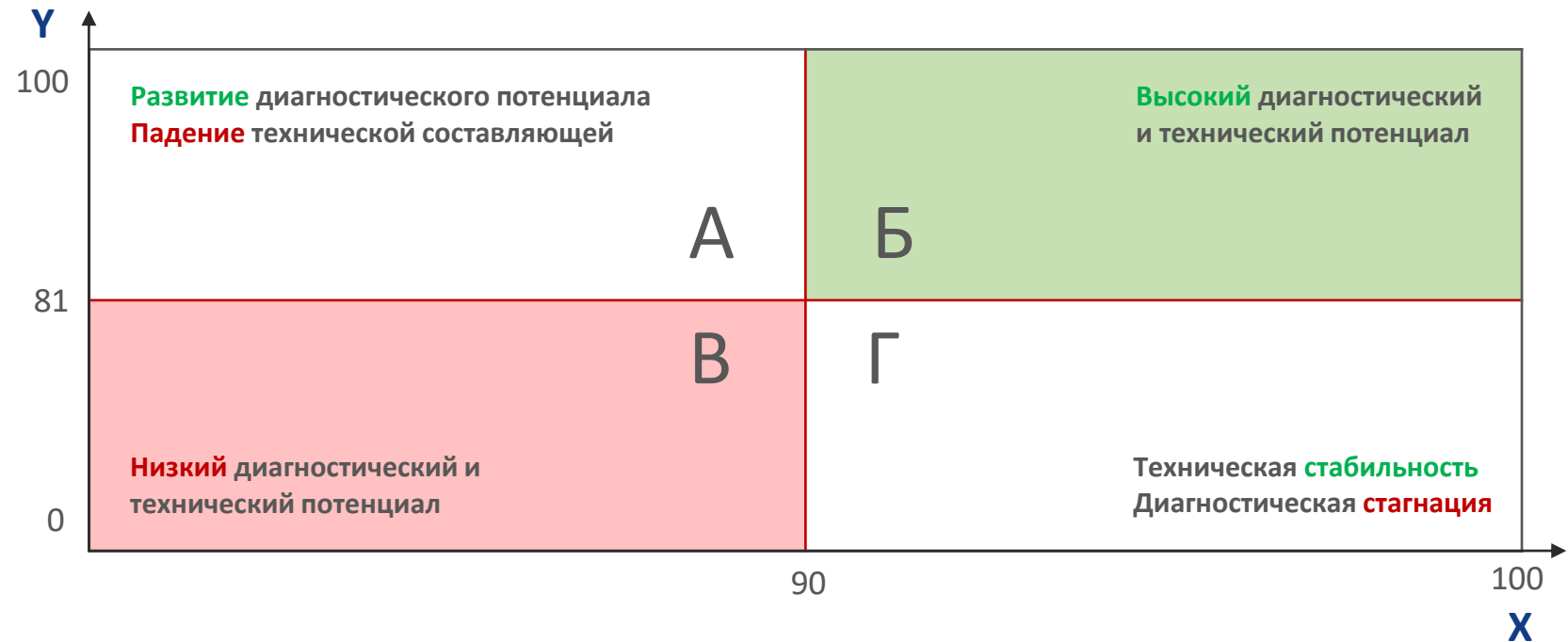
Ось 0Y – клиническая оценка
(проспективный ROC AUC * 100)

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



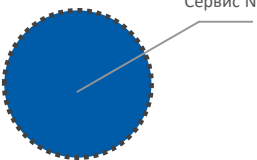
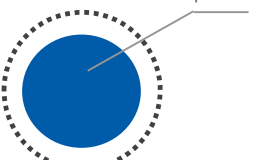
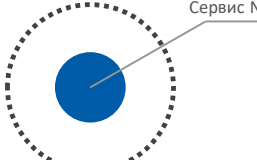
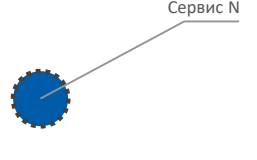
Ось 0X – техническая оценка (100 минус % технологических дефектов)

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



Эффективная составляющая матрицы



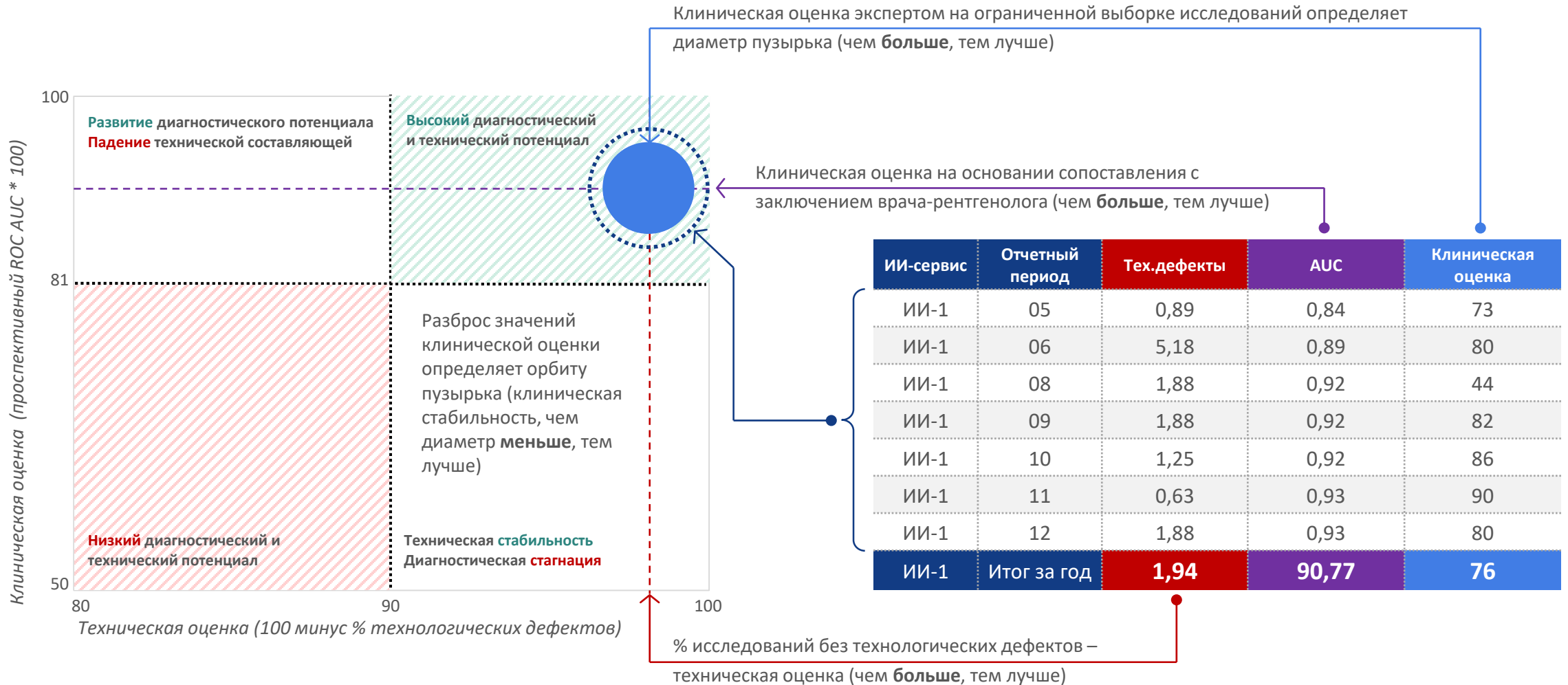
Результат работы ИИ				
Варианты эффективной составляющей	Пригодный и оптимальный результат	Пригодный и неоптимальный результат	Непригодный и неоптимальный результат	Непригодный и оптимальный результат

Диаметр пузырька – клиническая оценка работы ИИ-сервиса (на основании оценки корректности заключения и точности оконтуривания)

Орбита пузырька – разброс данных относительно **среднего** числа по клинической оценке (на основании серии клинических мониторингов)



Матрица. Пример.





ЦЕНТР ДИАГНОСТИКИ
И ТЕЛЕМЕДИЦИНЫ



БЛАГОДАРИМ ЗА ВНИМАНИЕ

✉ ai@zdrav.mos.ru

☎ +7 (495) 276 - 04 - 36

🌐 mosmed.ai