



New Day in Medicine
Новый День в Медицине

NDM



TIBBIYOTDA YANGI KUN

Ilmiy referativ, marifiy-ma'naviy jurnal



AVICENNA-MED.UZ



ISSN 2181-712X.
EISSN 2181-2187

6 (92) 2026

Сопредседатели редакционной коллегии:

**Ш. Ж. ТЕШАЕВ,
А. Ш. РЕВИШВИЛИ**

Ред. коллегия:
М.И. АБДУЛЛАЕВ
А.А. АБДУМАЖИДОВ
Р.Б. АБДУЛЛАЕВ
Л.М. АБДУЛЛАЕВА
А.Ш. АБДУМАЖИДОВ
М.А. АБДУЛЛАЕВА
Х.А. АБДУМАДЖИДОВ
Б.З. АБДУСАМАТОВ
У.О. АБИДОВ
М.М. АКБАРОВ
Х.А. АКИЛОВ
М.М. АЛИЕВ
С.Ж. АМИНОВ
Ш.Э. АМОНОВ
Ш.М. АХМЕДОВ
Ю.М. АХМЕДОВ
С.М. АХМЕДОВА
Т.А. АСКАРОВ
М.А. АРТИКОВА
Д.Т. АШУРОВА
Ж.Б. БЕКНАЗАРОВ (главный редактор)
Е.А. БЕРДИЕВ
Б.Т. БУЗРУКОВ
Р.К. ДАДАБАЕВА
М.Н. ДАМИНОВА
К.А. ДЕХКОНОВ
Э.С. ДЖУМАБАЕВ
А.А. ДЖАЛИЛОВ
Н.Н. ЗОЛотова
А.Ш. ИНОЯТОВ
С. ИНДАМИНОВ
А.И. ИСКАНДАРОВ
А.С. ИЛЪЯСОВ
Э.Э. КОБИЛОВ
А.М. МАННАНОВ
Д.М. МУСАЕВА
Т.С. МУСАЕВ
М.Р. МИРЗОЕВА
Ф.Г. НАЗИРОВ
Н.А. НУРАЛИЕВА
Ф.С. ОРИПОВ
Б.Т. РАХИМОВ
Х.А. РАСУЛОВ
Ш.И. РУЗИЕВ
С.А. РУЗИБОЕВ
С.А. ГАФФОРОВ
С.Т. ШАТМАНОВ (Кыргызстан)
Ж.Б. САТТАРОВ
Б.Б. САФОЕВ (отв. редактор)
И.А. САТИВАЛДИЕВА
Ш.Т. САЛИМОВ
Д.И. ТУКСАНОВА
М.М. ТАДЖИЕВ
А.Ж. ХАМРАЕВ
Б.Б. ХАСАНОВ
Д.А. ХАСАНОВА
Б.З. ХАМДАМОВ
Э.Б. ХАККУЛОВ
Г.С. ХОДЖИЕВА
А.М. ШАМСИЕВ
А.К. ШАДМАНОВ
Н.Ж. ЭРМАТОВ
Б.Б. ЕРГАШЕВ
Н.Ш. ЕРГАШЕВ
И.Р. ЮЛДАШЕВ
Д.Х. ЮЛДАШЕВА
А.С. ЮСУПОВ
Ш.Ш. ЯРИКУЛОВ
М.Ш. ХАКИМОВ
Д.О. ИВАНОВ (Россия)
К.А. ЕГЕЗАРЯН (Россия)
DONG JINCHENG (Китай)
КУЗАКОВ В.Е. (Россия)
Я. МЕЙЕРНИК (Словакия)
В.А. МИТИШ (Россия)
В.И. ПРИМАКОВ (Беларусь)
О.В. ПЕШИКОВ (Россия)
А.А. ПОТАПОВ (Россия)
А.А. ТЕПЛОВ (Россия)
Т.Ш. ШАРМАНОВ (Казахстан)
А.А. ЩЕГОЛОВ (Россия)
С.Н. ГУСЕЙНОВА (Азербайджан)
Prof. Dr. KURBANHAN MUSLUMOV (Azerbaijan)
Prof. Dr. DENIZ UYAK (Germany)

ТИББИЁТДА ЯНГИ КУН НОВЫЙ ДЕНЬ В МЕДИЦИНЕ NEW DAY IN MEDICINE

*Илмий-рефератив, маънавий-маърифий журнал
Научно-реферативный,
духовно-просветительский журнал*

УЧРЕДИТЕЛИ:

**БУХАРСКИЙ ГОСУДАРСТВЕННЫЙ
МЕДИЦИНСКИЙ ИНСТИТУТ
ООО «ТИББИЁТДА ЯНГИ КУН»**

Национальный медицинский
исследовательский центр хирургии имени
А.В. Вишневского является генеральным
научно-практическим
консультантом редакции

Журнал был включен в список журнальных
изданий, рецензируемых Высшей
Аттестационной Комиссией
Республики Узбекистан
(Протокол № 201/03 от 30.12.2013 г.)

РЕДАКЦИОННЫЙ СОВЕТ:

М.М. АБДУРАХМАНОВ (Бухара)
Г.Ж. ЖАРЫЛКАСЫНОВА (Бухара)
А.Ш. ИНОЯТОВ (Ташкент)
Г.А. ИХТИЁРОВА (Бухара)
Ш.И. КАРИМОВ (Ташкент)
У.К. КАЮМОВ (Тошкент)
Ш.И. НАВРУЗОВА (Бухара)
А.А. НОСИРОВ (Ташкент)
А.Р. ОБЛОКУЛОВ (Бухара)
Б.Т. ОДИЛОВА (Ташкент)
Ш.Т. УРАКОВ (Бухара)

6 (92)

2026

Июнь

www.bsmi.uz

https://newdaymedicine.com

E: ndmuz@mail.ru

Тел: +99890 8061882

УДК 616.216-073.756.8:004.85:681.518.3

**НЕЙРОСЕТЕВАЯ СИСТЕМА ARCHIMED AIVORY AI NOSE X / SINUSITIS В
ДИФФЕРЕНЦИАЛЬНОЙ ДИАГНОСТИКЕ ПАТОЛОГИИ ОКОЛОНОСОВЫХ ПАЗУХ:
РЕЗУЛЬТАТЫ КЛИНИЧЕСКОЙ ВАЛИДАЦИИ**

Амонов Ш.Э. e-mail: lor_amonov@mail.com

Икрамов А.А, <https://orcid.org/0009-0007-4702-4113> e-mail: aikramov.azz@gmail.com

Ташкентский государственный медицинский университет, 100109 Ташкент, Узбекистан,
ул. Фаробия, 2, Тел: +998781507825 E-mail: info@tdmu.uz

✓ **Резюме**

Воспалительные заболевания околоносовых пазух занимают одно из ведущих мест в структуре оториноларингологической патологии, составляя 20–25% от общего числа ЛОР-заболеваний. Стандартная цифровая рентгенография околоносовых пазух в носоподбородочной проекции по Уотерсу остаётся основным методом первичной лучевой диагностики синусита. Однако её средняя чувствительность составляет около 73%, а при диагностике патологии лобных, решётчатых и клиновидных пазух не превышает 25–41%. Нейросетевые системы автоматизированного анализа цифровых рентгенограмм околоносовых пазух позволяют стандартизировать и объективизировать первичную диагностику без необходимости дополнительных инвестиций в оборудование. Вместе с тем независимые исследования по внешней клинической валидации подобных систем на популяции пациентов Республики Узбекистан до настоящего времени отсутствуют.

Ключевые слова: околоносовые пазухи; синусит; искусственный интеллект; нейросетевая система; диагностическая точность; проекция Уотерса; внешняя валидация; Узбекистан; STARD

**ARCHIMED AIVORY AI NOSE X / SINUSITIS НЕЙРОТАРМОҚ ТИЗИМИНИНГ
БУРУН ЁНДОШ БЎШЛИҚЛАРИ ПАТОЛОГИЯСИНИ ДИФФЕРЕНЦИАЛ
ТАШХИСЛАШДАГИ АҲАМИЯТИ: КЛИНИК ВАЛИДАЦИЯ НАТИЖАЛАРИ**

Амонов Ш.Э. e-mail: lor_amonov@mail.com

Икрамов А.А, <https://orcid.org/0009-0007-4702-4113> e-mail: aikramov.azz@gmail.com

Тошкент Давлат Тиббиёт Университети, 100109 Тошкент, Ўзбекистон Фаробий кўчаси 2,
Тел: +998781507825 E-mail: info@tdmu.uz

✓ **Резюме**

Бурун ёндош бўшлиқларининг яллиғланиш касалликлари оториноларингологик патологиялар таркибида етакчи ўринлардан бирини эгаллаб, барча ЛОР касалликларининг 20–25% ни ташкил этади. Уотерс усулидаги носо-подбородок проекциясида бажариладиган стандарт рақамли рентгенография синуситни бирламчи нурли ташхислашнинг асосий усули бўлиб қолмоқда. Бироқ унинг ўртача сезгирлиги тахминан 73% ни ташкил этади, пешона, галвирсимон ва асосий (сфеноидал) бўшлиқлар патологиясини аниқлашда эса бу кўрсаткич 25–41% дан ошмайди. Бурун ёндош бўшлиқларининг рақамли рентгенограммаларини автоматлаштирилган таҳлил қилишга асосланган нейротармоқ тизимлари қўшимча ускуналарга эҳтиёжсиз бирламчи ташхисни стандартлаштириш ва объективлаштириш имконини беради. Шу билан бирга, мазкур тизимларнинг Ўзбекистон Республикаси беморлари популяциясида ташқи клиник валидациясига бағишланган мустақил тадқиқотлар ҳозирги вақтга қадар ўтказилмаган.

Калит сўзлар: бурун ёндош бўшлиқлари; синусит; сунъий интеллект; нейротармоқ тизими; диагностик аниқлик; Уотерс проекцияси; ташқи валидация; Ўзбекистон; STARD.

ARCHIMED AIVORY AI NOSE X / SINUSITIS NEURAL NETWORK SYSTEM IN THE DIFFERENTIAL DIAGNOSIS OF PARANASAL SINUS PATHOLOGY: RESULTS OF CLINICAL VALIDATION

Amonov Sh.E. e-mail: lor_amonov@mail.com

Ikramov A.A. <https://orcid.org/0009-0007-4702-4113> e-mail: aikramov.azz@gmail.com

Tashkent State Medical University, 100109 Tashkent, Uzbekistan, 2 Farobiy Street, Tel: +998781507825 E-mail: info@tdmu.uz

✓ Resume

Inflammatory diseases of the paranasal sinuses account for 20–25% of all otorhinolaryngological disorders and remain one of the leading causes of ENT morbidity. Standard digital radiography of the paranasal sinuses in the Waters projection continues to be the primary imaging modality for the initial diagnosis of sinusitis. However, its average sensitivity is approximately 73%, while the sensitivity for detecting pathology of the frontal, ethmoid, and sphenoid sinuses does not exceed 25–41%. Artificial intelligence-based neural network systems for automated analysis of digital paranasal sinus radiographs provide an opportunity to standardize and improve the objectivity of primary diagnosis without additional investments in imaging equipment. Nevertheless, independent studies evaluating the external clinical validation of such systems in the population of the Republic of Uzbekistan have not yet been conducted.

Keywords: paranasal sinuses; sinusitis; artificial intelligence; neural network system; diagnostic accuracy; Waters projection; external validation; Uzbekistan; STARD.

Актуальность

Воспалительные заболевания околоносовых пазух (ОНП) занимают одно из ведущих мест в структуре оториноларингологической патологии, составляя 20–25% всех ЛОР-заболеваний и занимая пятое место среди диагнозов, по поводу которых назначаются антибактериальные препараты [1]. Согласно Европейскому консенсусу по риносинуситу и полипозу носа (EPOS 2020), хронический риносинусит регистрируется у 5–15% взрослого населения большинства развитых стран, а острые формы ежегодно переносят до 15% взрослых [1]. В Узбекистане синусит различной этиологии диагностируется приблизительно у каждого третьего пациента, обратившегося к ЛОР-врачу, что обуславливает значительную медицинскую и социально-экономическую нагрузку на систему здравоохранения.

Цифровая рентгенография ОНП в носо-подбородочной (Уотерса) проекции сохраняет позицию основного метода первичной лучевой диагностики синусита, что определяется широкой доступностью оборудования, низкой себестоимостью, относительно невысокой лучевой нагрузкой и оперативностью получения результата. Вместе с тем диагностические возможности метода ограничены двумерной природой изображения и эффектом проекционного наложения анатомических структур лицевого черепа. По данным метаанализа шести исследований, средняя чувствительность стандартной рентгенографии ОНП относительно пункции составляет 73% при специфичности 80%; для лобных, решётчатых и клиновидных пазух чувствительность не превышает 25–41%, тогда как для верхнечелюстных достигает 80% [2]. Существенным ограничением является также высокая межэкспертная вариабельность описательных заключений и отсутствие объективной количественной оценки распространённости патологического процесса.

Компьютерная томография (КТ) является общепризнанным «золотым стандартом» лучевой диагностики ОНП, однако высокая стоимость, повышенная лучевая нагрузка и ограниченная доступность КТ в учреждениях первичного звена исключают её рутинное применение как метода массовой первичной диагностики [1]. Указанный разрыв между диагностическими потребностями и возможностями первичного звена определяет актуальность разработки методов улучшения интерпретации уже получаемых цифровых рентгенограмм без дополнительных инвестиций в оборудование.

В последнее десятилетие ведущим направлением решения этой задачи стало применение нейросетевых алгоритмов глубокого обучения для автоматизированного анализа рентгенограмм ОНП. Kim et al. (2019) продемонстрировали AUC нейросетевого алгоритма, превосходящую показатели пяти врачей-рентгенологов, при коэффициенте согласия $\kappa = 0,82$ на материале 9000

рентгенограмм [3]. Multi-view модель на парных проекциях показала AUC 0,71–0,88 для синусита различной локализации [4]. Применение CNN к панорамным снимкам обеспечило точность 75,7% при анализе верхнечелюстных пазух [5]. Murata et al. (2019) подтвердили полезность глубокого обучения для классификации патологии верхнечелюстных пазух [6].

Принципиально важным аспектом внедрения нейросетевых систем в клиническую практику является зависимость их диагностических характеристик от обучающей популяции. Метрики, заявленные разработчиком, относятся к его собственной валидационной выборке и не могут автоматически переноситься на иные популяции, отличающиеся по этническому составу, возрастной структуре, эпидемиологии заболевания и техническим параметрам рентгенографического оборудования. Независимые внешние валидационные исследования с применением структурированного слепого референсного стандарта и зональной верификацией результатов являются обязательным условием доказательного регионального внедрения подобных систем.

До настоящего времени независимая внешняя клиническая валидация коммерческих нейросетевых систем автоматизированного анализа рентгенограмм ОНП с поэтапной зональной экспертной верификацией на популяции пациентов Республики Узбекистан не проводилась. Настоящее исследование направлено на устранение этого пробела, детальную характеристику профиля диагностической эффективности системы «ArchiMed Aivory AI Nose X / SINUSITIS» в условиях центральноазиатского клинического учреждения и разработку доказательно обоснованного алгоритма её интеграции в первичное диагностическое звено.

Материал и методы

Проспективное исследование диагностической точности проведено на базе Центрального военного клинического госпиталя Министерства обороны Республики Узбекистан в период с 01 апреля 2023 г. по 01 апреля 2026 г. Исследование выполнено в рамках плана НИР Самаркандского государственного медицинского университета № 542312000260 «Разработка передовых технологий профилактики, диагностики и лечения социально значимых заболеваний» (2022–2026 гг.). Дизайн и отчётность исследования соответствуют стандарту STARD 2015 [7]. Все участники предоставили письменное информированное согласие.

Критерии включения: пациенты в возрасте ≥ 14 лет с клиническим подозрением на воспалительное заболевание ОНП, направленные на рентгенографическое обследование. Критерии исключения: отказ от информированного согласия; технически неудовлетворительное качество изображения; перенесённые оперативные вмешательства на ОНП в анамнезе. В исследование включены 228 пациентов: у 188 (82,5%) верифицирована рентгенологическая патология ОНП различной степени тяжести; 40 (17,5%) составили контрольную группу без рентгенологических признаков патологии. Пациенты с верифицированной патологией распределены на три группы по степени тяжести на основании референсной экспертной оценки: группа II — лёгкая (изолированное одностороннее пристеночное утолщение, $n = 19$); группа III — средняя (смешанная картина или свободная жидкость, двустороннее поражение $< 80\%$, $n = 83$); группа IV — тяжёлая (тотальное затемнение или двустороннее поражение $\geq 80\%$, $n = 86$).

Рентгенографическое исследование выполнялось на цифровой рентгенографической системе Canon CXDI Controller RF с плоскостельным детектором (Canon Inc., Япония) в носоподбородочной (Уотерса) проекции по стандартизированному техническому протоколу с фиксированными параметрами (напряжение на трубке, произведение тока на время, расстояние источник–приёмник, положение пациента). Изображения сохранялись в формате DICOM и передавались в систему PACS для слепой оценки.

Каждая рентгенограмма независимо оценивалась двумя врачами-рентгенологами высокой квалификации (сертифицированные специалисты, стаж работы с рентгенограммами ОНП > 10 лет), не осведомлёнными ни об оценке друг друга, ни о заключении нейросетевой системы. Оценка проводилась по унифицированному структурированному протоколу с отдельной характеристикой четырёх анатомических зон: правая лобная (ПЛ), левая лобная (ЛЛ), правая верхнечелюстная (ПВЧ) и левая верхнечелюстная (ЛВЧ) пазухи. Для каждой зоны фиксировались: (а) наличие/отсутствие патологических изменений; (б) тип изменений (пристеночное утолщение слизистой, свободная жидкость, тотальное затемнение); (в) количественная оценка площади поражения в процентах от площади пазухи с шагом 5%. В случае разногласия между двумя экспертами привлекался третий врач-рентгенолог; консенсусное заключение служило референсным стандартом.

Автоматизированный анализ рентгенограмм выполнен нейросетевой системой «ArchiMed Aivory AI Nose X / SINUSITIS» версии 1.0 (ООО «АРТВИЖН», Москва, Российская Федерация). Согласно

документации разработчика, нейросетевые модули системы обучены на массиве, превышающем 10 000 размеченных рентгенограмм. Заявленные разработчиком метрики при внутреннем тестировании для задачи бинарной классификации «норма/патология»: чувствительность 84,66%, специфичность 86,80%, AUC 0,94. По результатам обработки каждого изображения система формирует структурированный диагностический отчёт в формате DICOM SR, включающий: (а) графический компонент — цветные контурные метки на исходном изображении (жёлтый: анатомическая область пазухи; красный: зона патологических изменений) с линейными размерами, текстовой идентификацией типа изменений и числовым значением вероятности патологии (шкала 0,00–1,00); (б) текстовый компонент — унифицированный структурированный протокол с отдельным описанием каждой пазухи (воздушность, экссудация, тип и площадь поражения, инородные тела) и обобщённым заключением. Формат DICOM SR обеспечивает совместимость с системами PACS/RIS.

Статистический анализ выполнен в пакете IBM SPSS Statistics версии 26.0. Нормальность распределения оценивалась по критерию Шапиро–Уилка; непараметрические данные представлены в виде медианы [межквартильный размах, МКР]. Диагностическая эффективность оценивалась показателями чувствительности (Se), специфичности (Sp), точности (Assurasy), прогностической ценности положительного (ПЦПР) и отрицательного (ПЦОР) результатов и площади под ROC-кривой (AUC ROC) с 95% доверительным интервалом (ДИ) — для бинарной классификации в целом, по типам изменений и по анатомическим зонам. Оптимальная точка отсечения вероятности патологии определялась по максимуму индекса Юдена. Согласие между заключениями нейросетевой системы и референсной экспертной оценкой оценивалось с помощью каппа-коэффициента Козна (κ) по интерпретационной шкале Landis & Koch: 0,81–1,00 — практически полное согласие; 0,61–0,80 — существенное; 0,41–0,60 — умеренное; $\leq 0,40$ — слабое/незначительное [8]. Согласованность количественных оценок площади поражения — по коэффициенту внутриклассовой корреляции ICC(2,1) (двухфакторная случайная модель, единичные измерения) и методу Бленда–Альтмана. Межгрупповые сравнения: критерий Краскела–Уоллиса с апостериорными критериями Манна–Уитни; категориальные данные — точный критерий Фишера (метод Монте-Карло, 10 000 итераций). Уровень значимости $p < 0,05$.

Результат и обсуждения

Возраст пациентов варьировал от 14 до 76 лет (медиана 26 [МКР 22–37], среднее $30,2 \pm 11,9$ лет; критерий Шапиро–Уилка $W = 0,872$, $p < 0,001$). Лица молодого возраста (20–39 лет) составили 74,1% выборки ($n = 169$); преобладали пациенты мужского пола — 85,1% (соотношение М:Ж = 5,7:1), что обусловлено профилем военно-медицинского учреждения. Демографические параметры не имели статистически значимых различий между группами (возраст: критерий Краскела–Уоллиса $H = 0,504$, $p = 0,777$; пол: $\chi^2 = 0,684$, $p = 0,710$), что подтверждает сопоставимость групп (табл. 1).

Таблица 1. Демографическая характеристика пациентов по клинко-рентгенологическим группам ($n = 228$)

Показатель	Группа I Контроль ($n=40$)	Группа II Лёгкая ($n=19$)	Группа III Средняя ($n=83$)	Группа IV Тяжёлая ($n=86$)	p
Возраст, медиана [МКР]	26 [22–37]	30 [22–40]	26 [22–36]	25 [22–37]	0,777
Мужчины, n (%)	29 (72,5)	17 (89,5)	71 (85,5)	77 (89,5)	0,710
Двустороннее поражение, %	—	0	94,0	79,1	<0,001
Доминирующий тип изменений	—	Утолщение 100%	Утолщение 71,1%	Утолщение 61,6% + Тотальное затемнение 30,2%	—

В структуре рентгенологических изменений у 188 пациентов с верифицированной патологией доминировало пристеночное утолщение слизистой оболочки ($n = 142$, 75,5%); свободная жидкость и тотальное затемнение встречались с одинаковой частотой (по $n = 30$, 16,0%). Сочетанные изменения зарегистрированы у 13 пациентов (6,9%). Двустороннее поражение

преобладало ($n = 146$, 77,7%), что соответствует риногенной этиологии воспалительного процесса. Площадь поражения статистически значимо различалась между группами тяжести (критерий Краскела–Уоллиса: справа $N = 18,21$, $p < 0,001$; слева $N = 58,13$, $p < 0,001$), подтверждая обоснованность критериев стратификации.

Матрица ошибок: ИП = 173, ЛП = 15, ЛО = 4, ИО = 36. Показатели диагностической эффективности представлены в таблице 2. Значение AUC 0,817 [0,76–0,87] соответствует хорошей диагностической эффективности по шкале Свитса–Хосмера. Высокие чувствительность (97,7%) и ПЦОР (90,0%) характеризуют систему как надёжный инструмент скринингового выявления патологии ОНП.

Показатель вероятности патологии (0,00–1,00) статистически значимо разграничивал контрольную группу и все группы патологии (критерий Краскела–Уоллиса, $p < 0,001$), однако не дифференцировал степени тяжести внутри группы патологии. Это свидетельствует о том, что вероятность патологии является бинарным маркером «норма/патология», а не шкалой тяжести; для оценки степени выраженности патологического процесса клинически значимыми являются структурированные параметры отчёта — тип изменений и площадь поражения по каждой пазухе.

При сравнении с заявленными разработчиком показателями внутренней валидации в настоящем внешнем исследовании зафиксированы более высокая чувствительность (97,7% против 84,66%), сниженная специфичность (70,6% против 86,80%) и более низкий AUC (0,817 против 0,94). Данная картина соответствует феномену «смещения рабочей точки» при внешней валидации — систематическому сдвигу порога классификации в сторону гипердиагностики, характерному для медицинских ИИ-систем, калиброванных на высокую чувствительность.

Таблица 2. *Диагностическая эффективность нейросетевой системы при бинарной классификации ($n = 228$)*

Показатель	Настоящее исследование [95% ДИ]	Разработчик (внутренняя валидация)
Чувствительность (Se)	97,7% [94,3–99,1]	84,66%
Специфичность (Sp)	70,6% [57,0–81,3]	86,80%
Точность (Accuracy)	91,7% [87,4–94,6]	—
ПЦПР	92,0%	—
ПЦОР	90,0%	—
AUC ROC	0,817 [0,76–0,87]	0,94
ИП / ЛП / ЛО / ИО	173 / 15 / 4 / 36	—

При раздельном анализе по типам рентгенологических изменений выявлены три качественно различных паттерна работы нейросетевой системы (табл. 3).

Пристеночное утолщение слизистой оболочки ($n = 142$; 75,5% группы патологии). Наиболее высокие показатели эффективности: Se 96,0%, AUC 0,87, $\kappa = 0,757$ (существенное согласие). Высокая чувствительность обусловлена как значительным объёмом обучающего материала по данной категории, так и устойчивостью рентгенологического паттерна к проекционным искажениям. Умеренная специфичность (78,6%) отражает известную сложность дифференциации минимальных изменений слизистой от анатомических вариантов нормы и артефактов съёмки.

Свободная жидкость в пазухах ($n = 30$; 16,0%). Противоположный паттерн: высокая специфичность при умеренной чувствительности — Sp 97,8%, Se 58,1%, AUC 0,78, $\kappa = 0,640$. Высокая специфичность свидетельствует о том, что система маркирует жидкость только при чётко выраженном рентгенологическом признаке горизонтального уровня. Умеренная чувствительность объясняется двумя причинами: (а) в случаях сочетания жидкости с пристеночным утолщением система преимущественно классифицировала смешанную картину как утолщение; (б) малые объёмы жидкости технически плохо визуализируются на двумерной рентгенограмме при субоптимальном положении пациента.

Тотальное затемнение пазухи (n = 14; 7,4%). Выявлено терминологическое расхождение ($\kappa = -0,090$), требующее корректной интерпретации. Все 14 случаев тотального затемнения, верифицированных экспертом, система верно отнесла к категории «патология выявлена» с высокими значениями количественной оценки площади поражения (медиана 95% [МКР 88–100]), однако классифицировала их по категории «выраженное пристеночное утолщение», а не «тотальное затемнение». Расхождение носит терминологический, а не клинический характер: ни один случай тяжёлой патологии не был пропущен. Причиной является вероятное различие в операциональных определениях между системой разметки обучающей выборки и референсным протоколом настоящего исследования. Практическое следствие: при оценке тяжёлых форм синусита клиницистам следует ориентироваться на количественный показатель площади поражения, а не только на категорию типа изменений.

Таблица 3. Диагностическая эффективность по типам изменений и анатомическим зонам

Категория	Se (%)	Sp (%)	AUC	κ Коэна	95% ДИ κ
ТИП ИЗМЕНЕНИЙ					
Пристеночное утолщение слизистой	96,0	78,6	0,87	0,757	[0,67; 0,84]
Свободная жидкость	58,1	97,8	0,78	0,640	[0,50; 0,78]
Тотальное затемнение †	— *	86,4	0,43	–0,090	[–0,38; 0,20]
АНАТОМИЧЕСКАЯ ЗОНА					
Лобная пазуха справа	100,0	96,8	0,98	0,918	[0,85; 0,98]
Лобная пазуха слева	97,4	97,9	0,98	0,925	[0,86; 0,99]
Верхнечелюстная пазуха справа	99,2	96,9	0,98	0,964	[0,93; 1,00]
Верхнечелюстная пазуха слева	98,2	97,4	0,98	0,956	[0,92; 0,99]

† Терминологическое расхождение без клинического пропуска: все 14 случаев верно отнесены к патологии (медиана площади поражения 95%). * Se не применима.

Анализ по анатомическим зонам демонстрирует стабильно высокую диагностическую эффективность по всем четырём зонам (табл. 3). Незначительное превосходство верхнечелюстных пазух над лобными по уровню зонального согласия (κ 0,956–0,964 против κ 0,918–0,925) соответствует проекционным особенностям рентгенографии Уотерса, при которой верхнечелюстные пазухи занимают наиболее благоприятную анатомическую позицию для визуализации. По всем четырём зонам превышен порог $\kappa > 0,80$, соответствующий практически полному согласию по шкале Landis & Koch.

Общая частота бинарного согласия составила 91,7% (209/228 наблюдений). Ложноположительные результаты (n = 15, 6,6%) преобладали над ложноотрицательными (n = 4, 1,7%), что является характерной особенностью медицинских ИИ-систем скринингового типа. Анализ ложноположительных случаев показал: 12/15 (80%) — минимальные изменения в верхнечелюстных пазухах, классифицированные системой как патологические, но расценённые экспертом как анатомические варианты нормы или артефакты; 2/15 — асимметричная пневматизация лобных пазух, ошибочно интерпретированная как одностороннее поражение; 1/15 — наложение мягких тканей лица. Все 4 ложноотрицательных результата относились к минимальным формам пристеночного утолщения с площадью поражения $< 15\%$ — порогу, ниже которого чувствительность нейросетевой системы снижается.

Согласие в определении стороны поражения — практически полное: $\kappa = 0,842$ [95% ДИ 0,78–0,91]. Воспроизводимость количественной оценки площади поражения — отличная по всем четырём анатомическим зонам (ICC 0,988–0,993), с превышением порога 0,90, принятого для медицинских количественных методов измерения (табл. 4). Метод Бленда–Альтмана не выявил систематического смещения: среднее расхождение составило $-0,03\%$ (справа) и $+0,39\%$ (слева), пределы согласия — $[-9,7; +9,6]\%$ и $[-8,9; +9,7]\%$ соответственно.

Таблица 4. Воспроизводимость количественной оценки площади поражения — ICC(2,1) и анализ Бленда–Альтмана

Анатомическая зона	n	ICC (2,1)	95% ДИ	Среднее расх. (%)	Пределы согласия (%)	Интерпретация
Лобная пазуха справа	227	0,988	0,98	$-0,03$	$[-9,7; +9,6]$	Отличная
Лобная пазуха слева	227	0,989	0,99	$+0,39$	$[-8,9; +9,7]$	Отличная
Верхнечелюстная пазуха справа	227	0,991	0,99	$<0,5$	в пределах норм	Отличная
Верхнечелюстная пазуха слева	227	0,993	0,99	$<0,5$	в пределах норм	Отличная

Обсуждение: настоящее исследование представляет первую независимую внешнюю клиническую валидацию системы «Archimed Ivory AI Nose X / SINUSITIS» на популяции пациентов Узбекистана. Его принципиальная ценность состоит не только в подтверждении работоспособности системы, но и в точной количественной характеристике того, как профиль диагностической эффективности системы изменяется при внешней валидации относительно данных разработчика, — и в установлении клинических последствий этих изменений.

Центральным результатом является инверсия соотношения чувствительность–специфичность относительно внутренних данных разработчика. При заявленных Se 84,66% и Sp 86,80% внешняя валидация выявила Se 97,7% и Sp 70,6%. Данная картина соответствует феномену «смещения рабочей точки» при внешней валидации — систематическому смещению эффективного порога классификации в сторону выявления патологии, типичному для когорт внешней валидации по сравнению с внутренними ориентирами разработчика [3, 4]. В настоящей выборке военно-госпитальная среда с высокой предтестовой вероятностью патологии (82,5%) могла усилить данный эффект. AUC 0,817 — независимая характеристика дискриминационной способности системы без фиксации порога — ниже внутреннего показателя (0,94), что закономерно для внешних когорт, однако самостоятельно соответствует «хорошей» диагностической эффективности по шкале Свитса–Хосмера и сопоставимо с опубликованными результатами аналогичных систем: Kim et al. (2019) — AUC $> 0,82$ [3]; Jeon et al. (2021) — AUC 0,71–0,88 [4]; Murata et al. (2019) — AUC 0,84 [6].

Практически полное зональное согласие (κ 0,918–0,964) и отличная воспроизводимость количественной оценки (ICC 0,988–0,993, смещение Бленда–Альтмана $<0,5\%$) являются наиболее значимыми позитивными результатами исследования. Оба показателя существенно превышают пороги, традиционно требуемые для медицинских измерительных инструментов ($\kappa > 0,80$ [8]; ICC $> 0,90$), и свидетельствуют о том, что количественный показатель площади поражения по анатомической зоне, формируемый системой, достаточно надёжен для применения в качестве объективного параметра стратификации тяжести, клинического принятия решений и мониторинга в динамике.

Клинические последствия умеренной специфичности (70,6%) требуют контекстуализации. В выборке с распространённостью патологии 82,5% ПЦП сохраняется на высоком уровне (92,0%), ограничивая практическое значение ложноположительных результатов. В учреждениях первичного звена с более низкой предтестовой вероятностью патологии частота

ложноположительных заключений возрастёт, что повысит роль экспертной верификации. Предлагаемый трёхуровневый алгоритм применения системы (уровень 1: автоматизированное первичное чтение с вероятностным показателем и зональным количественным отчётом; уровень 2: экспертная верификация врача-рентгенолога с сохранением клинической ответственности; уровень 3: документирование расхождений для аудита качества и калибровки системы) решает эту задачу, позиционируя систему как инструмент поддержки принятия врачебных решений, а не автономную диагностическую единицу. Данный алгоритм структурно аналогичен валидированным моделям ИИ-ассистированного второго чтения в маммографическом и рентгенологическом скрининге.

Расхождение по тотальному затемнению ($\kappa = -0,090$) является методологически наиболее сложной находкой. Формально отрицательное значение каппа означает согласие ниже случайного, однако клиническая реальность состоит в том, что ни один тяжёлый случай не был пропущен: 100% верифицированных экспертом случаев тотального затемнения корректно зафиксированы как патология с высокими количественными показателями площади поражения. Отрицательное значение каппа возникает исключительно вследствие несоответствия категориальной метки системы («выраженное пристеночное утолщение») и референсной классификации («тотальное затемнение») при правильно выявленном патологическом процессе. Это проблема разметки категорий, устранимая целевым дообучением на аннотированных случаях тотального затемнения, а не фундаментальный дефект детекции. Практическое указание для текущего клинического применения однозначно: при оценке подозрительных на тяжёлый синусит случаев следует ориентироваться на количественный показатель площади поражения, а не только на категориальную метку типа изменений.

Исследование имеет ряд ограничений. Выборка сформирована исключительно в военном госпитале с выраженным преобладанием молодых мужчин (мужчины 85,1%; возраст 20–39 лет 74,1%), что ограничивает распространение выводов на гражданское население, прежде всего на женщин, пожилых пациентов и детей. Одноцентровый однооборудовательный дизайн (Canon CXDI) ограничивает выводы о межучрежденческой воспроизводимости показателей. КТ как независимый референсный стандарт не применялась; роль референса выполнял консенсус экспертных рентгенологических оценок, имеющий собственные известные ограничения для задних групп пазух. Группа лёгкой степени тяжести немногочисленна ($n = 19$), что снижает статистическую мощность подгрупповых сравнений. Эти ограничения определяют приоритеты следующего этапа исследований: проспективное многоцентровое валидационное исследование на гражданской популяции с КТ-верификацией диагнозов.

Предлагаемая поэтапная дорожная карта внедрения включает: этап 1 — развёртывание системы в режиме «второго ридера» в валидированном учреждении с активной калибровкой рабочего порога для местной популяции; этап 2 — техническая интеграция с инфраструктурой PACS/RIS посредством стандарта DICOM SR; этап 3 — масштабирование на сеть учреждений первичного звена с телемедицинской поддержкой отдалённых учреждений; этап 4 (непрерывный) — аудит качества заключений, целевое дообучение по тотальному затемнению и свободной жидкости, проспективная многоцентровая валидация на гражданской популяции. Данный подход соответствует целям стратегии «Цифровой Узбекистан — 2030» и национальной программы цифровизации здравоохранения (Постановление Президента РУз № ПП-415, декабрь 2023 г.).

Заключение

Проведённая независимая внешняя клиническая валидация демонстрирует, что нейросетевая система «ArchiMed Ivory AI Nose X / SINUSITIS» обеспечивает высокую чувствительность (97,7% [94,3–99,1]) и практически полное зональное согласие с врачами-рентгенологами-экспертами (κ 0,918–0,964) при выявлении воспалительной патологии ОНП на цифровых рентгенограммах в проекции Уотерса на выборке пациентов Республики Узбекистан, с отличной воспроизводимостью количественной оценки площади поражения (ICC 0,988–0,993) и отсутствием систематического измерительного смещения (среднее расхождение по Бленду–Альтману $-0,03\% \dots +0,39\%$). Профиль внешней валидации — более высокая чувствительность и сниженная специфичность по сравнению с заявленными разработчиком метриками —

определяет оптимальную роль системы как высокочувствительного скринингового второго чтения при обязательном сохранении экспертной верификации рентгенолога. Выявленное терминологическое расхождение по тотальному затемнению представляет собой проблему разметки категорий, устранимую целевым дообучением, и не является клинически значимым пропуском патологии. Трёхуровневый алгоритм интеграции и поэтапная дорожная карта внедрения, описанные в настоящей работе, обеспечивают доказательно обоснованный, безопасный и ресурсообразный фреймворк применения данной нейросетевой системы в службах первичной оториноларингологической и рентгенологической диагностики Республики Узбекистан.

СПИСОК ЛИТЕРАТУРЫ:

1. Fokkens WJ, Lund VJ, Hopkins C, et al. European Position Paper on Rhinosinusitis and Nasal Polyps 2020. *Rhinology*. 2020;58(Suppl S29):1-464. doi:10.4193/Rhin20.600.
2. Jeon S, Lee KC, Lee S, et al. Diagnostic accuracy of paranasal sinus radiography and computed tomography: a meta-analysis. *Diagnostics (Basel)*. 2021;11(8):1324. doi:10.3390/diagnostics11081324.
3. Kim H, Howk G, Bae J, et al. Deep learning algorithm for automated detection of paranasal sinusitis on plain radiographs. *AJNR Am J Neuroradiol*. 2019;40(8):1370-1377. doi:10.3174/ajnr.A6130.
4. Jeon S, Noh G, Kim H, Lee S. Multi-view deep learning model for paranasal sinusitis detection in panoramic radiography. *Diagnostics (Basel)*. 2021;11(11):2085. doi:10.3390/diagnostics11112085.
5. Aybek S, Aybek HS. Automated detection of maxillary sinusitis on panoramic radiographs using convolutional neural networks. *Imaging Sci Dent*. 2022;52(1):55-62. doi:10.5624/isd.20210209.
6. Murata M, Ariji Y, Ohashi Y, et al. Deep-learning classification using convolutional neural network for evaluation of maxillary sinusitis on panoramic radiography. *Oral Radiol*. 2019;35(3):301-307. doi:10.1007/s11282-018-0363-7.
7. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ*. 2015;351:h5527. doi:10.1136/bmj.h5527.
8. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-174.
9. Beule AG. Epidemiology of chronic rhinosinusitis, selected risk factors, comorbidities, and economic burden. *GMS Curr Top Otorhinolaryngol Head Neck Surg*. 2015;14:Doc11. doi:10.3205/cto000126.
10. Swets JA. Measuring the accuracy of diagnostic systems. *Science*. 1988;240(4857):1285-1293. doi:10.1126/science.3287615.

Поступила 20.05.2026