

## КАК ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ ВЛИЯЕТ НА ОЦЕНКУ ПОРАЖЕНИЯ ЛЁГКИХ ПРИ COVID-19 ПО ДАННЫМ КТ ГРУДНОЙ КЛЕТКИ?

© С.П. Морозов, В.Ю. Чернина, А.Е. Андрейченко, А.В. Владзимирский, О.А. Мокшенко, В.А. Гомболевский

Научно-практический клинический центр диагностики и телемедицинских технологий Департамента здравоохранения города Москвы, Москва, Российская Федерация

### АННОТАЦИЯ

**ОБОСНОВАНИЕ.** В период пандемии компьютерная томография (КТ) является одним из ключевых инструментов оценки изменений в лёгких, связанных с COVID-19. Рентгенологи Москвы используют адаптированную шкалу «КТ 0–4» для визуальной оценки зависимости тяжести общего состояния от характера и выраженности рентгенологических признаков изменений в лёгких при COVID-19 по данным КТ. В большом потоке исследований врач может пропустить находку и ошибиться в оценке объёма поражения лёгких, поэтому применение сервисов искусственного интеллекта (ИИ) обосновано в амбулаторном здравоохранении в период пандемии.

**ЦЕЛЬ** — сравнить распределение категорий «КТ 0–4» в заключениях, сформированных рентгенологами с использованием ИИ-сервисов и без них.

**МАТЕРИАЛ И МЕТОДЫ.** Ретроспективное исследование, ClinicalTrials (NCT04489992). Проанализированы результаты первичных КТ с категориями КТ 0–4 в период с 08.04.2020 по 01.12.2020 и отдельно за ноябрь 2020 года (с 01.11.2020 по 01.12.2020) в амбулаторных медицинских организациях Департамента здравоохранения. КТ проводились на 48 компьютерных томографах по стандартным протоколам, результаты обрабатывались через Единый радиологический информационный сервис. В тестовую группу включены КТ, обработанные ИИ-сервисами, в контрольную — без обработки ИИ. В анализ включены 5 ИИ-сервисов: RADlogics COVID-19 (RADLogics, США); COVID-IRA (IRA labs, Россия); Care Mentor AI, COVID (CareMentor AI, Россия); Третье Мнение. КТ-COVID-19 (Третье мнение, Россия); COVID-MULTIVOX (Гаммамед, Россия). ИИ-сервисы кодированы случайным образом.

**РЕЗУЛЬТАТЫ.** Проанализированы результаты КТ 260 594 пациентов (соотношение мужчины/женщины — 44/56%, средний возраст 49,5 года). В тестовую группу включены 115 618 КТ, в контрольную — 144 976. В зависимости от конкретного ИИ-сервиса для разных подгрупп категорий КТ-0 выставлено от 2,3 до 18,5% меньше, категорий КТ 3–4 — от 4,7 до 27,6% меньше, КТ-4 — от 40 до 60% меньше, чем в контрольной группе ( $p < 0,0001$ ).

За ноябрь (с 01.11.2020 по 01.12.2020) проанализированы результаты КТ 41 386 пациентов (соотношение мужчины/женщины — 44/56%, средний возраст — 53,2 года). В тестовую группу включено 28 881 КТ, в контрольную — 12 505. В зависимости от конкретного ИИ-сервиса для разных подгрупп категорий КТ-0, КТ 3–4 и КТ-4 выставлено соответственно от 1 до 2,6, от 0,2 до 15,7 и на 25% меньше, чем в контрольной группе ( $p=0,001$ ).

**ЗАКЛЮЧЕНИЕ.** Применение ИИ-сервисов для первичных КТ в амбулаторных условиях приводит к уменьшению количества выставляемых категорий КТ-0 и КТ 3–4, способных влиять на тактику ведения пациентов с COVID-19.

**Ключевые слова:** COVID-19; внебольничная пневмония; компьютерная томография; искусственный интеллект.

Как цитировать

Морозов С.П., Чернина В.Ю., Андрейченко А.Е., Владзимирский А.В., Мокиенко О.А., Гомболевский В.А. Как искусственный интеллект влияет на оценку поражения лёгких при COVID-19 по данным КТ грудной клетки? // *Digital Diagnostics*. 2021; 2(1):XX–XX. DOI: <https://doi.org/10.17816/DD60040>

Рукопись получена: 04.02.2021 Рукопись одобрена: 06.04.2021 Опубликована: 10.04.2021

## HOW DOES ARTIFICIAL INTELLIGENCE AFFECT THE ASSESSMENT OF LUNG LESIONS IN COVID-19 FROM CHEST CT SCANS?

© Sergey P. Morozov, Valeria Yu. Chernina, Anna E. Andreychenko, Valeria Yu. Chernina, Anton V. Vladzimirskiy, Olesya A. Mokienko, Victor A. Gombolevskiy

Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies of Moscow Health Care Department, Moscow, Russian Federation

### ABSTRACT

**BACKGROUND:** In the pandemic, chest computed tomography (CT) was a key tool in assessing lung changes associated with COVID-19. Radiologists in Moscow used an adapted CT 0–4 scale to visually assess the relationship between the severity of the general condition and the nature and severity of radiological signs of lung changes in COVID-19 on CT scans. In a large flow radiologists can miss findings and be mistaken in assessing the extent of lung damage, so the use of AI services may be of benefit in outpatient health care during a pandemic.

**AIMS:** To compare the distribution of CT 0–4 grades reported by radiologists with the help of AI and those generated without AI.

**MATERIALS AND METHODS:** A retrospective study, ClinicalTrials NCT04489992. We analyzed initial CT scans with categories CT 0–4 from 08.04.2020 to 01.12.2020 and separately for November (from 01.11.2020 to 01.12.2020) in outpatient medical organizations of the Moscow Healthcare Department. CT scans were performed on 48 units using standard protocols and processed through the Unified Radiology Information Service. The test group included scans processed by AI, and the control group included scans without AI. Five AI services were included: RADlogics COVID-19 (RADLogics, USA); COVID-IRA (IRA labs, Russia); Care Mentor AI, COVID (CareMentor AI, Russia); Third Opinion. CT-COVID-19 (Third Opinion, Russia); COVID-MULTIVOX (Gammamed, Russia). All services received random codes.

**RESULTS:** We analyzed 260 594 patients (male/female, % — 44/56, mean age — 49.5 years). The test group included 115 618 scans, and the control included 144 976. There were 2.3–18.5% fewer CT-0 categories for different vendors than in control. Categories CT 3–4 were used 4.7–27.6% less, and CT-4 was used 40–60% less compared to control ( $p < 0.0001$ ). Using AI for CT in COVID-19 changed the CT 0–4 grade in 4–18.9% (depending on the vendor).

For November (from 01.11.2020 to 01.12.2020) 41 386 patients (male/female, % — 44/56, mean age 53.2 years) were analyzed. The test group included 28 881 scans, and the control included 12 505. There were 1–2.6% fewer CT-0 categories for the different vendors than in control. There were 0.2–15.7% more CT 3–4 grades and CT-4 categories were used 25% less compared to control ( $p < 0.001$ ). Using AI for CT in COVID-19 changed the CT 0–4 grade in 14.1%.

**CONCLUSIONS:** AI for primary chest CT in the outpatient setting decreases in the number of CT-0 and CT 3–4 grades, possibly influencing patient management in COVID-19.

**Keywords:** COVID-19; pneumonia; multidetector computed tomography; artificial intelligence.

To cite this article

Morozov SP, Chernina VYu, Andreychenko AE, Vladzimirskyy AV, Mokienko OA, Gombolevskiy VA. How does artificial intelligence affect the assessment of lung lesions in COVID-19 from chest CT scans? *Digital Diagnostics*. 2020; 2(1):XX–XX. DOI: <https://doi.org/10.17816/DD60040>

Received: 04.02.2021 Accepted: 06.04.2021 Published: 10.04.2021

## СПИСОК СОКРАЩЕНИЙ И АББРЕВИАТУР

Датасет (англ. Data set) — набор данных, коллекция логических записей

ЕРИС — Единая радиологическая информационная система

ИИ-сервис — программное обеспечение на основе искусственного интеллекта

КТ — компьютерная томография

КТ 0–4 — эмпирическая шкала для визуальной оценки зависимости тяжести общего состояния от характера и выраженности рентгенологических признаков изменений в лёгких при COVID-19 по данным компьютерной томографии, где КТ-0 — отсутствие признаков вирусной пневмонии; КТ-1 — лёгкая форма пневмонии с участками «матового стекла», выраженность патологических изменений менее 25%; КТ-2 — умеренная пневмония, поражено 25–50% лёгких; КТ-3 — среднетяжёлая пневмония, поражено 50–75% лёгких; КТ-4 — тяжёлая форма пневмонии, поражено >75% лёгких

DICOM SR (DICOM — Digital Imaging and Communications in Medicine; SR — Structured Reporting) — стандартизованный формат медицинского отраслевого стандарта создания, хранения, передачи и визуализации цифровых медицинских изображений и документов обследованных пациентов (структурированная отчётность)

## ОБОСНОВАНИЕ

В 2020 году пандемия COVID-19 бросила вызов системам здравоохранения всего мира, что подтолкнуло правительства разных стран к поиску новых решений в условиях ограниченных ресурсов. 27 августа 2020 года правительственная комиссия по цифровому развитию РФ утвердила паспорт федерального проекта «Искусственный интеллект» в рамках национальной программы «Цифровая экономика». Независимо от этого, ещё в 2019 году было подготовлено Постановление Правительства Москвы, за которым в начале 2020 года вышел Приказ Департамента здравоохранения Москвы о проведении эксперимента по использованию инновационных технологий в области компьютерного зрения для анализа медицинских изображений и дальнейшего применения в системе здравоохранения города Москвы (Эксперимент) [1].

В пандемию метод компьютерной томографии (КТ) используют в качестве ключевого инструмента для оценки изменений в лёгких, связанных с инфекцией [2]. В первые месяцы пандемии COVID-19 с незначительной частотой в рутинной клинической работе для оценки выраженности изменений применяли преимущественно полуколичественные шкалы [3–7]. Работа со шкалами, в основу которых были заложены результаты исследований вирусной пневмонии SARS, предполагает отдельный расчёт объёма и вида поражений для долей и сегментов лёгких с последующим суммированием результатов [8]. Была предложена визуальная оценка, основанная на определении примерного объёма уплотнённой лёгочной ткани в обоих лёгких без отдельных расчётов для сегментов и долей [9].

Рентгенологи Москвы использовали адаптированную шкалу «КТ 0–4» для визуальной оценки зависимости тяжести общего состояния от характера и выраженности рентгенологических признаков изменений в лёгких при COVID-19 по данным КТ. Процент поражения оценивался отдельно по каждому лёгкому, степень изменений — по лёгкому с наибольшим поражением (вне зависимости от наличия постоперационных изменений). Каждые 25% объёма поражений лёгкого повышают шкалу на одну категорию [10, 11]. Валидация предложенной методики

визуальной оценки «КТ 0–4» проведена путём прогнозирования летальных исходов у больных COVID-19 [12]. При переходе из одной категории «КТ 0–4» в следующую риск смерти увеличивается в среднем на 38% (95% доверительный интервал 17,1–62,6) [12].

В большом потоке исследований врач может пропускать находки и ошибаться в оценке объёма поражения лёгких [13].

В Эксперимент была добавлена задача по обработке данных КТ органов грудной клетки для диагностики COVID-19 с помощью программного обеспечения на основе искусственного интеллекта (ИИ-сервисы). ИИ-сервисы добавляли КТ-серию с сегментацией поражений лёгких, информацию об объёме поражений для каждого лёгкого и категорию по шкале «КТ 0–4».

**ЦЕЛЬ** — сравнить распределение категорий «КТ 0–4» в заключениях, сформированных рентгенологами с использованием ИИ-сервисов и без них.

## МЕТОДЫ

### *Дизайн исследования*

Эксперимент является ретроспективным исследованием, в его основе — исследование, зарегистрированное в ClinicalTrials (NCT04489992). Данные, проанализированные в ходе работы, были предоставлены экспертами Департамента информационных технологий города Москвы.

### *Критерии соответствия*

Для формирования группы КТ-исследований были использованы критерии включения и исключения.

#### *Критерии включения:*

- КТ-исследования органов грудной клетки мужчин и женщин, обратившихся за медицинской помощью при подозрении на COVID-19;
- возраст пациентов старше 18 лет;
- КТ-исследования органов грудной клетки проведены и интерпретированы врачами-рентгенологами в период с 08.04.2020 по 01.12.2020 в амбулаторных медицинских организациях;
- наличие информации о категории оценки «КТ 0–4» в протоколах медицинских заключений врачей-рентгенологов;
- формирование протоколов описания КТ проведено в Единой радиологической информационной системе (ЕРИС).

#### *Критерии исключения:*

- КТ-исследования, в заключении которых указаны прочие изменения, не связанные с вирусной пневмонией.

### *Условия проведения*

КТ-исследования проводились во всех медицинских организациях, подведомственных Департаменту здравоохранения города Москвы, выполняющих амбулаторную помощь взрослому населению. Амбулаторные медицинские организации на время пандемии были трансформированы в амбулаторные КТ-центры, обеспечивающие специальный круглосуточный режим работы. С учётом эпидемиологической обстановки можно принять, что половозрастное распределение пациентов, которым было проведено КТ-исследование, соответствует такому же распределению в г. Москве.

### *Продолжительность исследования*

Исследование проведено в период с 08.04.2020 по 01.12.2020. Дополнительно проведена оценка в ноябре 2020 года (с 01.11.2020 по 01.12.2020) (рис. 1).

### *Описание медицинского вмешательства*

За весь рассматриваемый период и отдельно за ноябрь 2020 года были сформированы тестовая и контрольная группы сравнения (рис. 2). В тестовую группу включены КТ, обработанные ИИ-сервисами, в контрольную — без обработки ИИ.

Перед включением ИИ-сервиса в Эксперимент каждый алгоритм проходил тестирование на специально подготовленном калибровочном датасете. Калибровочный датасет включал КТ-исследования пациентов с лабораторно верифицированным COVID-19 и оценкой врачей-экспертов. Критерием включения в ЕРИС ИИ-сервиса являлась точность алгоритма не ниже площади под ROC-кривой (ROC AUC) 0,81, согласно методическим рекомендациям по клиническим испытаниям программного обеспечения на основе интеллектуальных технологий [14].

Каждый ИИ-сервис добавлял в исследование новую серию КТ, обработанную алгоритмами ИИ, и информацию в формате DICOM SR. Дополнительная серия КТ основана на оригинальной серии КТ текущего исследования с добавлением к оригинальному изображению сегментации поражений лёгких, которые, по предположению ИИ, вызваны COVID-19. Разработчикам ИИ, участвующим в эксперименте, было рекомендовано дополнять серию КТ, присланную ИИ-сервисом, суммарной информацией о поражении лёгких и оценкой по шкале «КТ 0–4». В данных DICOM SR, доступных каждому рентгенологу в тестовой группе во время формирования заключения, содержались информация о конкретном ИИ-сервисе, инструкция по пользованию результатами обработки и автоматически сформированное заключение, включающее оценку степени тяжести согласно шкале «КТ 0–4» (см. рис. 2).

КТ-исследования проводились на 48 компьютерных томографах (Toshiba Aquilion 64, Canon Medical Systems, Япония; HiSpeed GE, США; Optima CT 660, GE, США; Somatom Emotion 16, Siemens, Германия; Somatom Sensation 40, Siemens, Германия) по стандартным протоколам сканирования органов грудной клетки, рекомендуемым производителями.

Сравнение проводилось между категориями по шкале «КТ 0–4» из заключений врачей-рентгенологов, имеющих доступ к результатам ИИ-сервиса и не имеющих такого доступа.

### *Основной исход исследования*

ИИ-сервисы включились в Эксперимент после прохождения качественных и количественных тестирований на подготовленных экспертами ГБУЗ «Научно-практический клинический центр диагностики и телемедицинских технологий Департамента здравоохранения города Москвы» (НПКЦ ДиТ ДЗМ) базах данных. С момента включения пяти ИИ-сервисов исследования распределялись случайно с учётом вычислительных возможностей разработчиков. Одно исследование могло быть обработано несколькими ИИ-сервисами. ИИ-результаты использовались только для исследовательских целей, а финальное решение о категории по шкале «КТ 0–4» принимал рентгенолог.

### *Анализ в подгруппах*

В тестовой группе были сформированы подгруппы, так как в Эксперименте участвовали разные ИИ-сервисы для диагностики изменений в лёгких при COVID-19 по данным КТ.

В Эксперименте участвовали 7 разных ИИ-сервисов для диагностики COVID-19: RADlogics COVID-19 (RADLogics, США); COVID-IRA (IRA labs, Россия); Care Mentor AI, COVID (CareMertor AI, Россия); Третье Мнение. КТ-COVID-19 (Третье мнение, Россия); Multivox COVID19 (Гаммамед, Россия); ИРМ (Россия); CVL (СиВиЖнЛаб, Россия), однако последние два не вошли в тестовые подгруппы в связи с небольшим количеством обработанных исследований за весь период исследования. Остальные 5 ИИ-сервисов были кодированы случайным образом: ИИ-А, ИИ-В, ИИ-С, ИИ-Д, ИИ-Е (см. рис. 2).

Оценка данных проводилась на основе заключений рентгенологов, сформированных в ЕРИС.

### *Этическая экспертиза*

Получено одобрение Независимого этического комитета Московского регионального отделения Российского общества рентгенологов и радиологов (протокол №2 (1-II-2020) от 20.02.2020).

### Статистический анализ

*Принципы расчёта размера выборки:* в статистический анализ были включены все имеющиеся валидные данные. Методы восстановления отсутствующих данных не применялись.

*Методы статистического анализа данных:* для представления результатов были использованы методы описательной статистики с указанием абсолютного числа ( $n$ ) и доли (%) наблюдений в каждой категории. Межгрупповое сравнение распределения частоты в разных категориях между контрольной группой и тестовыми подгруппами внутри каждого из 2 периодов было проведено с помощью критерия согласия Пирсона ( $\chi^2$ ). Уровнем статистической значимости считалось значение 0,05. Статистический анализ был проведён с помощью программы Stata 14.

## РЕЗУЛЬТАТЫ

### Объекты (участники) исследования

За весь период проанализированы результаты первичных КТ-исследований 260 594 человек (соотношение мужчин/женщин — 44/56%, возраст от 18 до 100 лет, (средний возраст 49,5 года), выполненных и интерпретированных в период с 08.04.2020 по 01.12.2020 в амбулаторных медицинских организациях Москвы, перепрофилированных на время пандемии в амбулаторные КТ-центры.

### Основные результаты исследования

В тестовую группу было включено 115 618 КТ-исследований (44,4% общей выборки), в контрольную — 144 976 (55,6%). Соотношение исследований в контрольной и тестовой группе было однородным. Распределение по подгруппам в тестовой группе: ИИ-А — 98 953 исследований (37,9% общей выборки), ИИ-В — 44 194 (17%), ИИ-С — 24 067 (9,2%), ИИ-Д — 22 679 (8,7%), ИИ-Е — 10 645 (4,1%).

Для разных подгрупп с ИИ-сервисами выставлено от 2,3 до 18,5% меньше категорий КТ-0 (поражений COVID-19 не выявлено), чем в контрольной группе. Для разных подгрупп с ИИ-сервисами выставлено от 4,7 до 27,6% меньше категорий КТ 3–4, чем в контрольной группе. Для разных подгрупп с ИИ-сервисами выставлено от 40 до 60% меньше категорий КТ-4, чем без ИИ (рис. 3;  $p < 0,0001$ ).

За ноябрь 2020 года проанализированы результаты первичных КТ-исследований 41 386 человек (соотношение мужчин/женщин — 44/56%, возраст от 18 до 100 лет, в среднем 53,2 года), выполненных и интерпретированных в период с 01.11.2020 по 01.12.2020 в амбулаторных медицинских организациях Москвы, перепрофилированных на время пандемии в амбулаторные КТ-центры.

В тестовую группу было включено 28 881 КТ-исследование (69,8% общей выборки за ноябрь), в контрольную — 12 505 (30,2%). Распределение по подгруппам в тестовой группе: ИИ-А — 12 266 исследований (29,6% общей выборки за ноябрь), ИИ-В — 11 922 (28,8%), ИИ-С — 9785 (23,6%), ИИ-Д — 9942 (24%), ИИ-Е — 8602 (20,8%).

Для разных подгрупп с ИИ-сервисами выставлено от 1 до 2,6% меньше категорий КТ-0 (поражений COVID-19 не выявлено), чем в контрольной группе. Для разных подгрупп с ИИ-сервисами выставлено от 0,2 до 15,7% больше категорий КТ 3–4, чем без ИИ. Для разных подгрупп с ИИ-сервисами выставлено на 25% меньше категорий КТ-4, чем без ИИ (рис. 4;  $p = 0,0010$ ).

Для периода «Ноябрь 2020» статистически значимые отличия были выявлены только для КТ-0, по остальным категориям отличий не было. Но даже для КТ-0 они минимальны — 18,6 против 17,0%. По общему периоду наблюдения, напротив, все категории статистически значимо отличаются между данными «без ИИ» и «суммарно по всем ИИ». Критическое значение  $\chi^2$

равно 4. Для суммарного периода все категории «КТ 0–4» вносят значительный вклад в различия. Минимальное значение  $\chi^2$  равно 26,2 для КТ-3 ( $p < 0,0001$ ).

## ОБСУЖДЕНИЕ

### *Резюме основного результата исследования*

В результатах исследования выявлено изменение между категориями по шкале «КТ 0–4», сформированными рентгенологами при наличии результатов обработки ИИ-сервисами, и категориями, сформированными без использования ИИ-сервисов.

### *Обсуждение основного результата исследования*

В целом, до пандемии COVID-19 ИИ-сервисы использовались при выявлении рентгенологических симптомов для обнаружения заболеваний, классификации, оптимизации изображений, снижения дозы лучевой нагрузки и улучшения рабочего процесса [15].

Медицинские исследования делают приложения ИИ более понятными, безопасными, эффективными и интегрированными в рабочие процессы врачей [16]. Важно понимать, что валидация алгоритма ИИ должна включать не только ретроспективный анализ точности по сравнению с золотым стандартом, но и рандомизированное клиническое исследование для оценки влияния ИИ на принятие решений медицинскими работниками [17]. Например, в рандомизированном клиническом исследовании NURE показан эффект системы ранней диагностики интраоперационной гипотензии на основе машинного обучения: в группе интервенции медианное время гипотензии составило 8,0 мин против 32,7 в контрольной группе ( $p < 0,001$ ), число летальных исходов в группе интервенции составило 0 против 2 в контрольной группе [18]. В другом крупном исследовании с применением ИИ для анализа рентгенограмм органов грудной клетки при COVID-19 20% опрошенных врачей сообщили, что алгоритм повлиял на принятие клинических решений [19].

В доступной литературе не обнаружено подобных исследований, оценивающих влияние результатов ИИ на принятие окончательного решения рентгенологами во время диагностики изменений в лёгких при подозрении на COVID-19 по данным КТ грудной клетки.

Перед ИИ для COVID-19 стоят две задачи — выявление и классификация. Задача выявления определяется разницей между КТ-0 и всеми остальными категориями. Задача классификации заключается в выявлении отличий между разными категориями КТ (КТ1–КТ4), т.е. разными объёмами поражения лёгочной паренхимы.

Первой находкой является разница удельного веса категории КТ-0 между всеми подгруппами. В случае с ИИ-сервисом А врачу-рентгенологу демонстрировалась только часть срезов с наличием изменений, характерных для COVID-19, тогда как все остальные ИИ-сервисы имели такое же количество срезов, как и в оригинальной серии КТ. Мнения врачей без ИИ-А близки к контрольной группе.

Второй находкой является разница удельного веса категорий КТ 3–4 между всеми подгруппами. Вероятно, при большом количестве поражений лёгких эмпирический глазомер рентгенолога может преувеличивать объём поражения. Такая гипердиагностика компенсируется, когда рентгенолог проконтролирует сегментацию повреждённых лёгких, выполненную с помощью ИИ. Благодаря этому повышается степень доверия к автоматическому расчёту объёма и категории по шкале «КТ 0–4». Поскольку в качестве обоснования госпитализации использовался порог поражения лёгкого  $>50\%$  (категории КТ 3–4), в сочетании с клинико-лабораторными данными это может уменьшать количество госпитализаций.

На основании публикаций С.П. Морозова и соавт. по прогнозированию летальных исходов при COVID-19 по данным КТ органов грудной клетки, при переходе из одной категории КТ в следующую риск увеличивался в среднем на 38% (95% доверительный интервал 17,1–62,6), а в дополнительном исследовании среди пациентов с лабораторно верифицированным COVID-19 риск смерти с категорией КТ-4 был выше в 3 раза, чем при КТ-0 [12, 20]. В нашем исследовании

продемонстрировано, что категорий КТ-4 в тестовой подгруппе меньше, чем в контрольной. В предыдущих исследованиях показано, что оценка степени поражения лёгких, соответствующая КТ-4, связана с летальными исходами при COVID-19 больше, чем все остальные категории (КТ 0–3) [12, 20]. Именно поэтому изменение количества пациентов с категориями КТ-4 имеет важное значение для формирования калькуляторов рисков летальных исходов для пациентов с COVID-19.

Из-за пандемии одновременный запуск всех ИИ-сервисов был признан ограничивающим потенциальную пользу от применения ИИ, так как пришлось бы ожидать всех разработчиков для участия в Эксперименте, поэтому на протяжении 2020 года наблюдается неравномерность присоединения к Эксперименту разных ИИ-сервисов. Для получения дополнительных результатов был проведён анализ за ноябрь, где количество КТ-исследований в контрольной группе и каждой тестовой подгруппе сопоставимо между собой.

Выявленная разница между результатами первого и второго периодов, по мнению авторов, связана с несколькими факторами:

- 1) разное количество ИИ-сервисов;
- 2) технический фактор: до 2020 года ИИ-сервисы, участвующие в Эксперименте, не имели возможности обучать свои алгоритмы для оценки и диагностики поражения лёгких для диагностики COVID-19, поэтому в процессе Эксперимента была признана оправданной возможность изменения версионности программного обеспечения на основе ИИ в целях повышения качества работы алгоритмов и потенциально большей пользы;
- 3) человеческий фактор: до 2020 года врачи не использовали оценку КТ грудной клетки по шкале «КТ 0–4», которая представляет определённую сложность с точки зрения оценки объёма множественных поражений в лёгких. Следует отметить, что врачи могли самостоятельно повышать свои навыки по оценке объёма поражений лёгких при COVID-19, так как оценивали значительное количество КТ-исследований в период пандемии. Кроме этого, врачи могли набираться опыта, проверяя разметку, выполненную ИИ-сервисом, что могло привести к улучшению навыка в более правильной оценке объёма поражений.

В ЕРИС врач-рентгенолог имеет возможность оставить отзыв о работе ИИ-сервиса в специальном поле для обратной связи. Перспективой исследования является сравнение распределения категорий по шкале «КТ 0–4» у врачей-рентгенологов, не встречающихся с ИИ-сервисами на протяжении пандемии и пользующихся результатами ИИ-сервисов на основе их обратной связи.

### *Ограничения исследования*

Наше исследование имеет ряд ограничений. Не вошли в него пациенты с положительными результатами теста полимеразной цепной реакции для верификации COVID-19, так как результаты этих исследований после проведения КТ. Исследование не было рандомизированным. Не проводилась оценка степени согласия рентгенологов с результатами ИИ-сервисов. В тестовой группе часть КТ анализировались несколькими ИИ-сервисами. ИИ-сервисы не имели регистрации как медицинские изделия. В течение пандемии ИИ-сервисы менялись с повышением качества обработки КТ, и этот факт дополнительно не оценивался в данном исследовании. Не учитывалась адаптация рентгенологов к использованию шкалы «КТ 0–4».

В ЕРИС, где врачи формировали медицинские заключения, было встроено специально поле для обратной связи по вопросам работы ИИ-сервисов. Однако на момент формирования публикации проводится обработка результатов обратной связи от врачей, что не позволяет представить её в текущем исследовании.

## **ЗАКЛЮЧЕНИЕ**

Результаты показывают, что применение ИИ-сервисов для первичных КТ органов грудной клетки в амбулаторных условиях приводит к уменьшению количества категорий КТ-0 и КТ 3–4, что способно влиять на тактику ведения пациентов при COVID-19.

Необходимо дополнительное исследование, которое позволит оценить, насколько снижение выбора вышеперечисленных категорий целесообразно для ведения пациентов, как изменение маршрутизации в дальнейшем влияет на показатели выздоровления и смертности.

## ДОПОЛНИТЕЛЬНО

**Источник финансирования.** Исследование и публикация статьи осуществлены на личные средства авторского коллектива.

**Конфликт интересов.** Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

**Участие авторов.** Авторы подтверждают соответствие своего авторства международным критериям ICMJE (прочли и одобрили финальную версию перед публикацией). Наибольший вклад распределён следующим образом: С.П. Морозов — концепция исследования; В.Ю. Чернина — поиск публикаций по теме статьи, написание текста рукописи; А.Е. Андрейченко — формирование набора данных, редактирование текста рукописи; А.В. Владзимирский — редактирование текста рукописи; О.А. Мокиенко — экспертная оценка информации, редактирование текста рукописи; В.А. Гомболевский — концепция исследования, экспертная оценка информации, написание текста рукописи, утверждение финальной версии рукописи.

**Благодарности:** Авторы выражают глубокую признательность коллективам отделений лучевой диагностики медицинских организаций Департамента Здравоохранения Москвы. Авторы благодарят за сплоченную работу коллективов Департамента информационных технологий города Москвы и ООО «Лаваль», обеспечивающих внедрение искусственного интеллекта в практическое здравоохранение на уровне крупного мегаполиса. Кроме этого, авторы отдельно благодарят разработчиков компании Binomics ray, RADLogics, IRA labs, CareMentor AI, Третье мнение, Гаммамед. Каждый из участников процесса приложил важные усилия к исследованию в непростой период эпидемии. Авторы отдельно благодарят О.В. Омелянскую, Е.Г. Бахтееву, И.А. Виноградову, С.О. Ермолаева, Л.Г. Родионову, К.В. Хрипунову, К.М. Арзамасова, П.А. Николаева, С.Ф. Четверикова, И.А. Блохина за административно-организационную, педагогическую и тестовую работу при подготовке и реализации Эксперимента; отдельная благодарность - В.Г. Кляшторному за статистический анализ.

**Funding source.** The study had no sponsorship.

**Competing interests.** The authors declare that they have no competing interests.

**Author contribution.** The authors confirm that they meet the ICMJE international criteria for authorship (have read and approved the final version before publication). The largest contributions are distributed as follows: S.P. Morozov - research concept; V.Yu. Chernina - search for publications on the topic of the article, writing the text of the manuscript; A.E. Andreychenko - data set formation, editing the manuscript text; A.V. Vladzimirskyy - editing the manuscript text; O.A. Mokienko - expert information evaluation, editing the manuscript text; V.A. Gombolevskiy - research concept, expert information evaluation, writing the manuscript text, approval of the final manuscript version.

**Acknowledgements:** The authors would like to express their gratitude to the teams of radiology departments of medical organizations of the Department of Healthcare of Moscow. The authors would like to thank the teams of the Department of Information Technology of the City of Moscow and Laval LLC for their concerted efforts in implementing artificial intelligence in practical healthcare at the level of a large metropolitan city. In addition, the authors thank separately the developers of Binomics ray, RADLogics, IRA labs, CareMentor AI, Third Opinion and Gammamed. Each of the contributors made important research efforts during a difficult time of the epidemic. The authors thank separately O.V. Omelyanskaya, E.G. Bakhteeva, I.A. Vinogradova, S.O. Ermolaev, L.G. Rodionova, K.V. Khripunova, K.M. Arzamasov, P.A. Nikolaev, S.F. Chetverikov, I.A. Blokhin, for administrative,

pedagogical, and test work in preparing and implementing the Experiment; special thanks to V.G. Klyashorny for statistical analysis.

## СПИСОК ЛИТЕРАТУРЫ

1. Эксперимент по использованию инновационных технологий в области компьютерного зрения для анализа медицинских изображений и дальнейшего применения в системе здравоохранения города Москвы [дата обращения: 04.02.2021]. Режим доступа: <https://mosmed.ai>. Дата обращения: 15.01.2021.
2. Morozov S.P., Ledikhova N.V., Panina E.V., et al. Re: Controversy in coronaViral Imaging and Diagnostics (COVID) // Clin Radiol. 2020. Vol. 75, N 11. P. 871–872. doi: 10.1016/j.crad.2020.07.023
3. Chang Y.C., Yu C.J., Chang S.C., et al. Pulmonary sequelae in convalescent patients after severe acute respiratory syndrome: evaluation with thin-section CT // Radiology. 2005. Vol. 236, N 3. P. 1067–1075. doi: 10.1148/radiol.2363040958
4. Haseli S., Khalili N., Bakhshayeshkaram M., et al. Lobar distribution of COVID-19 pneumonia based on chest computed tomography findings. A retrospective study // Arch Acad Emerg Med. 2020. Vol. 8, N 1. P. 55.
5. Inui S., Fujikawa A., Jitsu M., et al. Chest CT findings in cases from the cruise ship «Diamond Princess» with Coronavirus Disease 2019 (COVID-19) // Radiol Cardiothorac Imaging. 2020. Vol. 2, N 2. P. 200110. doi: 10.1148/ryct.2020200110
6. Prokop M., van Everdingen W., van Rees Vellinga T., et al. CO-RADS: A Categorical CT assessment scheme for patients suspected of having COVID-19-definition and evaluation // Radiology. 2020. Vol. 296, N 2. P. 97–104. doi: 10.1148/radiol.2020201473
7. Shen C., Yu N., Cai S., et al. Quantitative computed tomography analysis for stratifying the severity of Coronavirus Disease 2019 // J Pharm Anal. 2020. Vol. 10, N 2. P. 123–129. doi: 10.1016/j.jpha.2020.03.004
8. Pan F., Ye T., Sun P., et al. Time course of lung changes at chest CT during Recovery from Coronavirus Disease 2019 (COVID-19) // Radiology. 2020. Vol. 295, N 3. P. 715–721. doi: 10.1148/radiol.2020200370
9. Revel M.P., Parkar A.P., Prosch H., et al. COVID-19 patients and the radiology department – advice from the European Society of Radiology (ESR) and the European Society of Thoracic Imaging (ESTI) // Eur Radiol. 2020. Vol. 30, N 9. P. 4903–4909. doi: 10.1007/s00330-020-06865-у
10. Морозов С.П., Проценко Д.Н., Сметанина С.В., и др. Лучевая диагностика коронавирусной болезни (COVID-19): организация, методология, интерпретация результатов: препринт № ЦДТ-2020-II. Версия 2 от 17.04.2020. Серия «Лучшие практики лучевой и инструментальной диагностики». Вып. 65. Москва: ГБУЗ «НПКЦ ДиТ ДЗМ», 2020. 78 с.
11. Сеницын В.Е., Тюрин И.Е., Митьков В.В. Временные методические рекомендации Российского общества рентгенологов и радиологов (РОРР) и Российской ассоциации специалистов ультразвуковой диагностики в медицине (РАСУДМ) «Методы лучевой диагностики пневмонии при новой коронавирусной инфекции при COVID-19» (версия 2) // Вестник рентгенологии и радиологии. 2020. Т. 101, № 2. С. 72–89. doi: 10.20862/0042-4676-2020-101-2-72-89
12. Морозов С.П., Гомболевский В.А., Чернина В.Ю., и др. Прогнозирование летальных исходов при COVID-19 по данным компьютерной томографии органов грудной клетки // Туберкулез и болезни лёгких. 2020. Т. 98, № 6. С. 7–14. doi: 10.21292/2075-1230-2020-98-6-7-14

13. Howard J. Cognitive errors and diagnostic mistakes. A case-based guide to critical thinking in medicine. New York: Springer; 2019.
14. Морозов С.П., Владимирский А.В., Кляшторный В.Г., и др. Клинические испытания программного обеспечения на основе интеллектуальных технологий (лучевая диагностика). Серия «Лучшие практики лучевой и инструментальной диагностики». Вып. 57. Москва: ГБУЗ «НПКЦ ДиТ ДЗМ», 2019. 51 с.
15. Sahiner B., Pezeshk A., Hadjiiski L.M., et al. Deep learning in medical imaging and radiation therapy // *Med Phys*. 2019. Vol. 46, N 1. P. 1–36. doi: 10.1002/mp.13264
16. Allen B.J., Seltzer S.E., Langlotz C.P., et al. A road map for translational research on artificial intelligence in medical imaging: from the 2018 national institutes of health/RSNA/ACR/The academy workshop // *J Am Coll Radiol*. 2019. Vol. 16, N 9. P. 1179–1189. doi: 10.1016/j.jacr.2019.04.014
17. Angus D.C. Randomized clinical trials of artificial intelligence // *Jama*. 2020. Vol. 323, N 11. P. 1043–1045. doi: 10.1001/jama.2020.1039
18. Wijnberge M., Geerts B.F., Hol L., et al. Effect of a machine learning-derived early warning system for intraoperative hypotension vs standard care on depth and duration of intraoperative hypotension during elective noncardiac surgery: the HYPE randomized clinical trial // *Jama*. 2020. Vol. 323, N 11. P. 1052–1060. doi: 10.1001/jama.2020.0592
19. Carlile M., Hurt B., Hsiao A., et al. Deployment of artificial intelligence for radiographic diagnosis of COVID-19 pneumonia in the emergency department // *J Am Coll Emerg Physicians Open*. 2020. Vol. 1, N 6. P. 1459–1464. doi: 10.1002/emp2.12297
20. Морозов С.П., Чернина В.Ю., Блохин И.А., Гомболевский В.А. Прогнозирование исходов при лабораторно верифицированном COVID-19 по данным компьютерной томографии органов грудной клетки: ретроспективный анализ 38051 пациента // *Digital Diagnostics*. 2020. Т. 1, № 1. С. 27–36. doi: 10.17816/DD46791

## REFERENCES

1. Experiment on the use of innovative computer vision technologies for medical image analysis and subsequent applicability in the healthcare system of Moscow [cited 2021 Feb 04]. (In Russ). Available from: <https://mosmed.ai>
2. Morozov SP, Ledikhova NV, Panina EV, et al. Re: Controversy in coronaViral Imaging and Diagnostics (COVID). *Clin Radiol*. 2020;75(11):871–872. doi: 10.1016/j.crad.2020.07.023
3. Chang YC, Yu CJ, Chang SC, et al. Pulmonary sequelae in convalescent patients after severe acute respiratory syndrome: evaluation with thin-section CT. *Radiology*. 2005;236(3):1067–1075. doi: 10.1148/radiol.2363040958
4. Haseli S, Khalili N, Bakhshayeshkaram M, et al. Lobar distribution of COVID-19 pneumonia based on chest computed tomography findings. A retrospective study. *Arch Acad Emerg Med*. 2020;8(1):e55.
5. Inui S, Fujikawa A, Jitsu M, et al. Chest CT findings in cases from the cruise ship "Diamond Princess" with Coronavirus Disease 2019 (COVID-19). *Radiol Cardiothorac Imaging*. 2020;2(2):e200110. doi: 10.1148/ryct.2020200110
6. Prokop M, van Everdingen W, van Rees Vellinga T, et al. CO-RADS: A Categorical CT assessment scheme for patients suspected of having COVID-19-definition and evaluation. *Radiology*. 2020;296(2):97–104. doi: 10.1148/radiol.2020201473
7. Shen C., Yu N., Cai S., et al. Quantitative computed tomography analysis for stratifying the severity of Coronavirus Disease 2019. *J Pharm Anal*. 2020;10(2):123–129. doi: 10.1016/j.jpha.2020.03.004
8. Pan F, Ye T, Sun P, et al. Time course of lung changes at chest CT during recovery from Coronavirus Disease 2019 (COVID-19). *Radiology*. 2020;295(3):715–721. doi: 10.1148/radiol.2020200370
9. Revel MP, Parkar AP, Prosch H, et al. COVID-19 patients and the radiology department – advice from the European Society of Radiology (ESR) and the European Society of Thoracic Imaging (ESTI). *Eur Radiol*. 2020;30(9):4903–4909. doi: 10.1007/s00330-020-06865-y

10. Morozov SP, Protsenko DN, Smetanina SV, et al. Radiation diagnosis of coronavirus disease (COVID-19): organization, methodology, interpretation of results: preprint II. Version 2 of 17.04.2020. The series "Best practices of radiation and instrumental diagnostics". Issue 65. Moscow: Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies of the Moscow Health Care Department; 2020. 78 p. (In Russ).
11. Sinitsyn VE, Tyurin IE, Mitkov VV. Consensus Guidelines of Russian Society of Radiology (RSR) and Russian Association of Specialists in Ultrasound Diagnostics in Medicine (RASUDM) "Role of Imaging (X-ray, CT and US) in Diagnosis of COVID-19 Pneumonia" (version 2). *Journal of Radiology and Nuclear Medicine*. 2020;101(2):72–89. (In Russ). doi: 10.20862/0042-4676-2020-101-2-72-89
12. Morozov SP, Gombolevskiy VA, Chernina VY, et al. Prediction of lethal outcomes in COVID-19 cases based on the results chest computed tomography. *Tuberculosis and Lung Diseases*. 2020;98(6):7–14. (In Russ). doi: 10.21292/2075-1230-2020-98-6-7-14
13. Howard J. Cognitive errors and diagnostic mistakes. A case-based guide to critical thinking in medicine. New York: Springer; 2019.
14. Morozov SP, Vladzimirsky AV, Klyashtorny VG, et al. Clinical trials of software based on intelligent technologies (radiation diagnostics). The series "Best practices of radiation and instrumental diagnostics". Issue 57. Moscow: Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies of the Moscow Health Care Department; 2019. 51 p. (In Russ).
15. Sahiner B, Pezeshk A, Hadjiiski LM, et al. Deep learning in medical imaging and radiation therapy. *Med Phys*. 2019;46(1):1–36. doi: 10.1002/mp.13264
16. Allen BJ, Seltzer SE, Langlotz CP, et al. A road map for translational research on artificial intelligence in medical imaging: from the 2018 national institutes of health/RSNA/ACR/The academy workshop. *J Am Coll Radiol*. 2019;16(9):1179–1189. doi: 10.1016/j.jacr.2019.04.014
17. Angus DC. Randomized clinical trials of artificial intelligence. *Jama*. 2020;323(11):1043–1045. doi: 10.1001/jama.2020.1039
18. Wijnberge M, Geerts BF, Hol L, et al. Effect of a machine learning-derived early warning system for intraoperative hypotension vs standard care on depth and duration of intraoperative hypotension during elective noncardiac surgery: the HYPE randomized clinical trial. *Jama*. 2020;323(11):1052–1060. doi: 10.1001/jama.2020.0592
19. Carlile M, Hurt B, Hsiao A, et al. Deployment of artificial intelligence for radiographic diagnosis of COVID-19 pneumonia in the emergency department. *J Am Coll Emerg Physicians Open*. 2020;1(6):1459–1464. doi: 10.1002/emp2.12297
20. Morozov SP, Chernina VYu, Blokhin IA, et al. Chest computed tomography for outcome prediction in laboratory-confirmed COVID-19: A retrospective analysis of 38,051 cases. *Digital Diagnostics*. 2020;1(1):27–36. (In Russ). doi: 10.17816/DD46791

ОБ АВТОРАХ	AUTHOR'S INFO
<p><b>*Гомболевский Виктор Александрович</b>, к.м.н.; адрес: Россия, 127051, Москва, ул. Петровка, д. 24/1; ORCID: <a href="https://orcid.org/0000-0003-1816-1315">https://orcid.org/0000-0003-1816-1315</a>; eLibrary SPIN: 6810-3279; e-mail: <a href="mailto:v.gombolevskiy@npcmr.ru">v.gombolevskiy@npcmr.ru</a></p>	<p><b>*Victor A. Gombolevskiy</b>, Cand. Sci. (Med.); address: 24/1 Petrovka, Moscow, 127051, Russia; ORCID: <a href="https://orcid.org/0000-0003-1816-1315">https://orcid.org/0000-0003-1816-1315</a>; eLibrary SPIN: 6810-3279; e-mail: <a href="mailto:v.gombolevskiy@npcmr.ru">v.gombolevskiy@npcmr.ru</a></p>
<p><b>Морозов Сергей Павлович</b>, д.м.н., профессор; ORCID: <a href="https://orcid.org/0000-0001-6545-6170">https://orcid.org/0000-0001-6545-6170</a>; eLibrary SPIN: 8542-1720; e-mail: <a href="mailto:morozov@npcmr.ru">morozov@npcmr.ru</a></p>	<p><b>Sergey P. Morozov</b>, Dr. Sci. (Med.), Professor; ORCID: <a href="https://orcid.org/0000-0001-6545-6170">https://orcid.org/0000-0001-6545-6170</a>; eLibrary SPIN: 8542-1720; e-mail: <a href="mailto:morozov@npcmr.ru">morozov@npcmr.ru</a></p>

<b>Андрейченко Анна Евгеньевна</b> , к.ф.-м.н.; ORCID: <a href="https://orcid.org/0000-0001-6359-0763">https://orcid.org/0000-0001-6359-0763</a> ; eLibrary SPIN: 6625-4186; e-mail: <a href="mailto:a.andreychenko@npcmr.ru">a.andreychenko@npcmr.ru</a>	<b>Anna E. Andreychenko</b> , Cand. Sci. (Phys.-Math.); ORCID: <a href="https://orcid.org/0000-0001-6359-0763">https://orcid.org/0000-0001-6359-0763</a> ; eLibrary SPIN: 6625-4186; e-mail: <a href="mailto:a.andreychenko@npcmr.ru">a.andreychenko@npcmr.ru</a>
<b>Чернина Валерия Юрьевна</b> ; ORCID: <a href="http://orcid.org/0000-0002-0302-293X">http://orcid.org/0000-0002-0302-293X</a> ; eLibrary SPIN: 8896-8051; e-mail: <a href="mailto:v.chernina@npcmr.ru">v.chernina@npcmr.ru</a>	<b>Valeria Yu. Chernina</b> , MD; ORCID: <a href="http://orcid.org/0000-0002-0302-293X">http://orcid.org/0000-0002-0302-293X</a> ; eLibrary SPIN: 8896-8051; e-mail: <a href="mailto:v.chernina@npcmr.ru">v.chernina@npcmr.ru</a>
<b>Владзимирский Антон Вячеславович</b> , д.м.н.; ORCID: <a href="https://orcid.org/0000-0002-2990-7736">https://orcid.org/0000-0002-2990-7736</a> ; eLibrary SPIN: 3602-7120; e-mail: <a href="mailto:a.vladzimirsky@npcmr.ru">a.vladzimirsky@npcmr.ru</a>	<b>Anton V. Vladzimirskyy</b> , Dr. Sci. (Med.); ORCID: <a href="https://orcid.org/0000-0002-2990-7736">https://orcid.org/0000-0002-2990-7736</a> ; eLibrary SPIN: 3602-7120; e-mail: <a href="mailto:a.vladzimirsky@npcmr.ru">a.vladzimirsky@npcmr.ru</a>
<b>Мокиенко Олеся Александровна</b> , к.м.н.; ORCID: <a href="https://orcid.org/0000-0002-7826-5135">https://orcid.org/0000-0002-7826-5135</a> ; eLibrary SPIN: 8088-9921; e-mail: <a href="mailto:Lesya.md@yandex.ru">Lesya.md@yandex.ru</a>	<b>Olesya A. Mokienko</b> , Cand. Sci. (Med.); ORCID: <a href="https://orcid.org/0000-0002-7826-5135">https://orcid.org/0000-0002-7826-5135</a> ; eLibrary SPIN: 8088-9921; e-mail: <a href="mailto:Lesya.md@yandex.ru">Lesya.md@yandex.ru</a>

ARTICLE IN PRESS



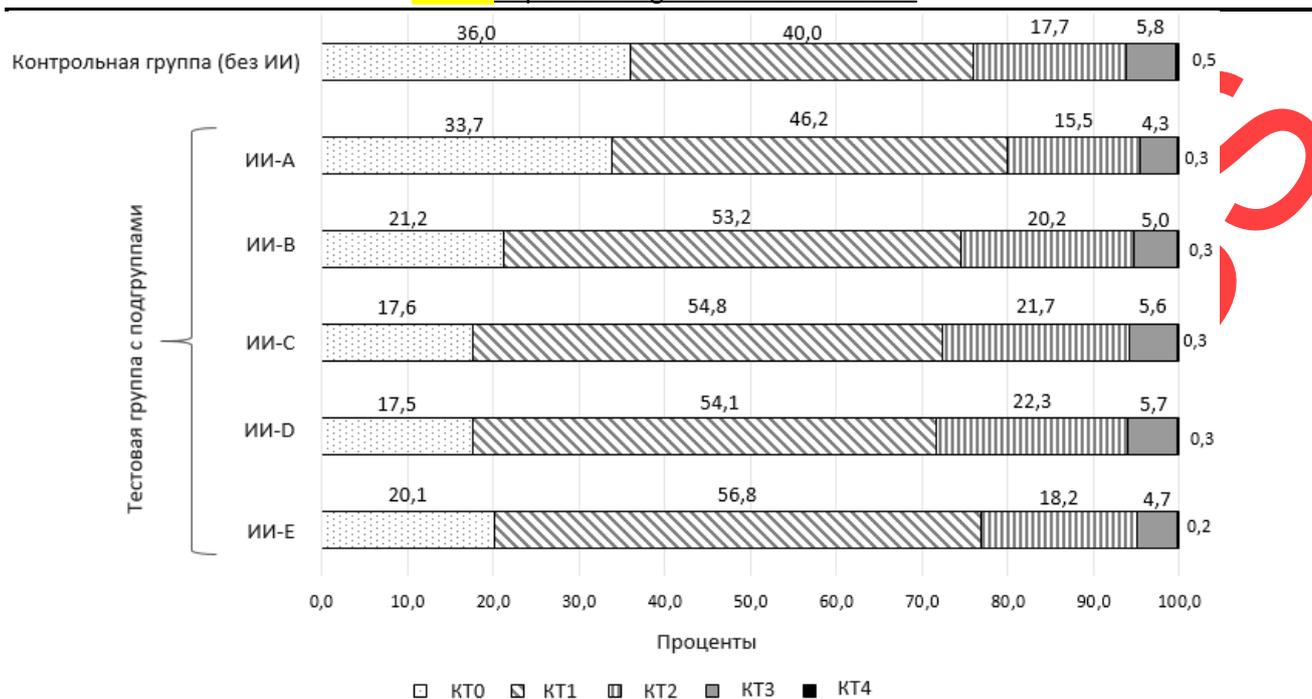


Рис. 3. Результаты сравнения первичных КТ органов грудной клетки, выполненных в амбулаторных КТ-центрах, по степени тяжести категорий «КТ 0–4» между контрольной группой и тестовыми подгруппами за весь период (08.04–01.12.2020).  
*n*=260 594; *p* < 0,0001.

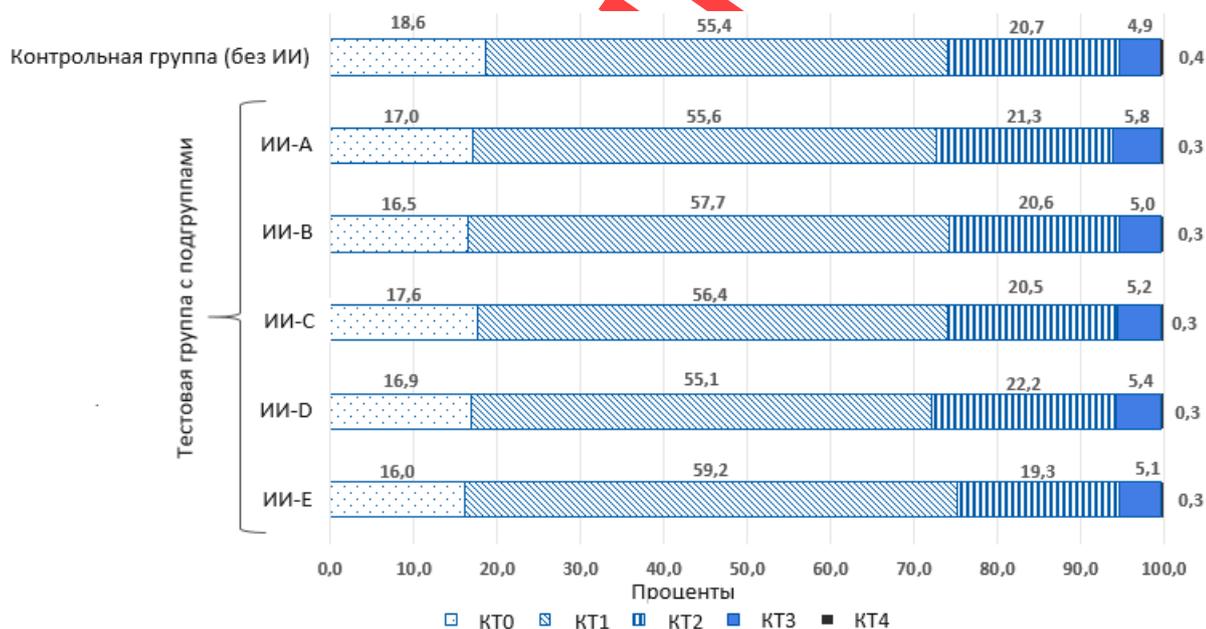


Рис. 4. Результаты сравнения первичных КТ органов грудной клетки, выполненных в амбулаторных КТ-центрах, по степени тяжести категорий «КТ 0–4» между контрольной группой и тестовыми подгруппами за ноябрь 2020 г.  
*n*=41 386; *p*=0,0010.