

DOI: <https://doi.org/10.17816/DD60040>



How does artificial intelligence effect on the assessment of lung damage in COVID-19 on chest CT scan?

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

Moscow Center for Diagnostics and Telemedicine, Moscow, Russian Federation

BACKGROUND: During the pandemic, computed tomography (CT) was one of the most important tools for assessing COVID-19-related lung changes. In COVID-19 patients, radiologists in Moscow used the adapted CT0-4 scale to visually assess the dependence of the severity of the general condition on the nature and severity of radiological signs of changes in the lungs based on computed tomography. In a large stream of scans, the doctor may miss findings and make errors in assessing the volume of lung damage, so the use of AI services in outpatient healthcare during a pandemic can be beneficial.

AIM: The goal of this study is to compare the distribution of CT0-4 categories designed by radiologists with the results of AI services processing and categories formed without AI services.

METHODS: We used retrospective study design, full study protocol is registered on ClinicalTrials.gov (NCT04489992). The results of primary CT scans with the CT0-4 categories were analyzed in outpatient medical institutions of the Health Department from April 08, 2020, to December 01, 2020, and separately for November (from November 01, 2020, to December 01, 2020). CT was performed on 48 computed tomographs in accordance with standard protocols, and the data was processed by the single radiology information systems. CTs in the test group received AI services, while CTs in the control group did not. The analysis includes five AI services: RADLogics COVID-19 (RADLogics, USA), COVID-IRA (IRA labs, Russia), Care Mentor AI, COVID (Care Mentor AI, Russia), Third Opinion. CT-COVID-19 (Third Opinion, Russia), and COVID-MULTIVOX (Gammamed, Russia). Moreover, AI services are encoded at random.

RESULTS: The CT scan results of 260,594 patients were examined (m/f % = 44/56, mean age = 49.5). The test group consisted of 115,618 CT scans, while the control group consisted of 144,976 CT scans. Depending on the specific AI service, CT0 was established by 2.3–18.5% less than the control group for different subgroups of categories. The categories CT3-4 were established by 4.7–27.6% less than without AI, and the categories CT4 by 40–60% less than without AI ($p < 0.0001$). For November (from November 01, 2020, to December 01, 2020), the CT scan results of 41,386 patients were analyzed (m/f % = 44/56, average age = 53.2 years). The test group consisted of 28,881 CT scans, while the control group included 12,505 CT scans. Depending on the specific AI service, CT0 was established by 1–2.6% less than the control group for different subgroups of categories. Further, the categories CT3–CT4 were established by 0.2–15.7% less than without AI, and the categories CT4 were established by 25% less than without AI ($p = 0.001$).

CONCLUSION: The use of AI services for primary CT scans on an outpatient basis reduces the number of CT0 and CT3–CT4 results, which can influence the therapeutic approach for COVID-19 patients.

Keywords: COVID-19; community-acquired pneumonia; computed tomography; artificial intelligence.

To cite this article

Morozov SP, Chernina VYu, Andreychenko AE, Vladzimirskyy AV, Mokienko OA, Gomboleviskiy VA. How does artificial intelligence effect on the assessment of lung damage in COVID-19 on chest CT scan? *Digital Diagnostics*. 2020;2(1):27–38. DOI: <https://doi.org/10.17816/DD60040>

DOI: <https://doi.org/10.17816/DD60040>

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. Мокиенко, В.А. Гомболевский

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

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, протокол исследования зарегистрирован в ClinicalTrials.gov (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. Т. 2, №1. С. 27–38. DOI: <https://doi.org/10.17816/DD60040>

DOI: <https://doi.org/10.17816/DD60040>

人工智能如何影响胸部CT扫描对COVID-19中肺损伤的评估？

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

Moscow Center for Diagnostics and Telemedicine, Moscow, Russian Federation

理由：在大流行期间，计算机断层扫描（CT）是评估与COVID-19相关的肺部变化的主要工具之一。莫斯科的放射学家使用了经过调整的KT0-4量表，根据计算机断层扫描技术，通过视觉评估了一般病情严重程度对COVID-19中肺部改变的放射学征象的性质和严重程度的依赖性。大量的研究中，医生可能会遗漏发现结果并在评估肺损伤量方面犯错误，因此在大流行期间，在门诊医疗中使用AI服务可能很有用。

目的：比较放射科医生形成的CT0-4类别的分布与AI服务处理的结果以及没有AI服务形成的类别的比较。

方法：回顾性研究，ClinicalTrials.gov (NCT04489992)。DZM的门诊医疗组织中，分析了从CT0-4类别进行的一次CT扫描的结果，分析时间为：2020年4月8日至2020年1月12日，以及11月（2020年11月1日至2020年1月12日）。根据标准协议在48台计算机断层扫描仪上执行CT，并通过ERIS处理。测试组包括由AI服务处理的CT，对照组为不包含AI的CT。分析包括5种AI服务：RADlogics COVID-19（美国RADLogics），COVID-IRA（俄罗斯的IRA实验室），Care Mentor AI，COVID（俄罗斯的CareMentor AI），第三意见。CT-COVID-19英寸（第三意见，俄罗斯），COVID-MULTIVOX（俄罗斯伽马迈德）。AI服务是随机编码的。

结果：分析了260594例患者的CT扫描结果（m / f% = 44/56，平均年龄-49.5）。测试组包括115,618次CT扫描，对照组-144976。根据特定的AI服务，对于CT-0类别的不同子组，其设置比对照组少2.3%至18.5%。与未使用AI相比，将CT3-4类别设置为比不使用AI少4.7%至27.6%，并且将CT-4类别与不使用AI设置成从40%至60%（p < 0.0001）。

对于11月（从01.11.2020到01.12.2020），分析了41386名患者的CT扫描结果（m / f% = 44/56，平均年龄-53.2岁）。测试组包括28881 CT扫描，对照组-12505。根据特定的AI服务，对于CT-0类别的不同子组，其设置比对照组小1%至2.6%。显示的CT3-4类别比没有使用AI的类别多出0.2%至15.7%；类别CT-4设置为比不使用AI时少25%（p = 0.001）。

结论：在门诊基础上将AI服务用于主要CT扫描会导致CT-0和CT3-4数量减少，从而影响管理COVID-19患者的策略。

关键词：COVID-19；社区获得性肺炎；CT扫描；人工智能。

引用本文：

Morozov SP, Chernina VYu, Andreychenko AE, Vladzimirskiy AV, Mokienko OA, Gombolevskiy VA. 人工智能如何影响胸部CT扫描对COVID-19中肺损伤的评估? *Digital Diagnostics*. 2020;2(1):27-38. DOI: <https://doi.org/10.17816/DD60040>

收到: 04.02.2021

接受: 06.04.2021

发布时间: 09.04.2021

Acronyms and abbreviations

data set: a set of data, a collection of logical records

URIS: Unified Radiological Information System

AI service: artificial intelligence software

CT: computed tomography

CT (computed tomography) 0–4 represents an empirical scale for visual assessment on the severity dependence of general condition of the nature and pronouncement of radiological changes in the lungs with Coronavirus disease 2019 (COVID-19) according to CT, where CT-0 implies the absence of signs of viral pneumonia; CT-1 indicates mild pneumonia with ground glass opacity areas,

with <25% severity of pathological changes; CT-2 is moderate pneumonia, with affection of 25%–50% of the lungs; CT-3 implies moderate pneumonia with 50%–75% of the lungs affected; and CT-4 is severe pneumonia, affecting >75% of the lungs.

Digital Imaging and Communications in Medicine: Structured Reporting is a standardized format of the medical industry standard for creating, storing, transferring, and visualizing digital medical images and documents of patients examined (structured reporting).

BACKGROUND

In 2020, the Coronavirus disease 2019 (COVID-19) pandemic challenged healthcare systems worldwide, which prompted global governments to seek new solutions under resource-constrained conditions. On August 27, 2020, the government commission on digital development of the Russian Federation approved the certificate of the federal project “Artificial Intelligence” within the national program “Digital Economy”. Regardless of this, back in 2019, a Decree of the Moscow Government was drafted, followed by an Order of the Moscow Healthcare Department at the beginning of 2020 on conducting an experiment on the use of innovative technologies in the field of computer vision for the analysis of medical images and further application in the health care system of Moscow (Experiment) [1].

In the pandemic, computed tomography (CT) is used as a key tool for assessing changes in the lungs, associated with infection [2]. In the first months of the COVID-19 pandemic, semi-quantitative scales were mainly used to assess the severity of changes, with an insignificant frequency in routine clinical work [3–7]. Work with scales, which were based on severe acute respiratory syndrome viral pneumonia studies results, involves a separate calculation of the volume and type of lesions for lobes and segments of the lungs with subsequent summation of results [8]. A visual assessment was proposed based on the approximate volume determination of indurated tissue in both lungs without separate calculations for segments and lobes [9].

Radiologists in Moscow used the adapted CT 0–4 scale to assess visually the severity dependence of the general condition on the nature and pronouncement of radiological changes in the lungs with COVID-19 according to CT data. The percentage of damage was assessed separately for each lung, and the degree of change was assessed for the lung with the greatest lesion (regardless of postoperative changes). Every 25% of the volume of lung lesions increases the scale by one category [10, 11]. The proposed method of visual assessment CT 0–4 was validated by predicting lethal

outcomes in patients with COVID-19 [12]. The risk of lethal outcome increases by an average of 38% (95% confidence interval 17.1–62.6) with the transition from one CT 0–4 category to the next [12].

In case of a large stream of studies, the doctor may omit findings and make mistakes in assessing lung damage volume [13].

A task has been added to the experiment to process chest CT data for COVID-19 diagnostics using artificial intelligence software (AI services). The AI services added CT series with lung lesion segmentation, lesion volume information for each lung, and a CT 0–4 category.

This study aimed to compare the distribution of CT 0–4 categories in reports generated by radiographers with and without AI services.

METHODS

Study design

This is a retrospective study based on a study registered in Clinical Trials (NCT04489992). Data analyzed in the course of the work was provided by experts of the Moscow Information Technology Department.

Compliance criteria

Inclusion and exclusion criteria were used to form the CT study group.

Inclusion criteria:

- CT scans of the chest organs of male and female patient who sought medical help with suspected COVID-19;
- age of patients over 18 years old;
- CT examinations of the chest organs were performed and interpreted by radiologists in the period from April 8, 2020 to December 1, 2020 in outpatient healthcare organizations;

- availability of information on the assessment category by CT 0–4 in protocols of medical reports of radiologists;
- CT description protocols were formed in the Unified Radiological Information System (URIS).

Exclusion criteria:

- CT studies with conclusion of other changes not associated with viral pneumonia.

Implementation conditions

CT examinations were conducted in all medical organizations subordinate to the Moscow Healthcare Department, providing outpatient care for the adult population. During the pandemic, outpatient medical organizations were transformed into outpatient CT centers that provided a special round-the-clock operation. Taking into account the epidemiological situation, it can be assumed that the gender and age distribution of patients who underwent CT scan corresponds to the same distribution in Moscow.

Study duration

The study was conducted in the period from April 8, 2020 to December 1, 2020. Additionally, an assessment was performed in November 2020 (from November 1, 2020 to December 1, 2020) (Fig. 1).

Description of the medical intervention

Over the entire period under consideration and separately for November 2020, the test and control comparison groups were formed (Fig. 2). The test group included CTs processed by AI services, and the control group included CTs without AI processing.

Each algorithm was tested on a specially prepared calibration data set before including the AI service in the experiment. The calibration data set included CT scans of patients with laboratory-verified COVID-19 and an assessment by expert doctors. The criterion for inclusion of an AI service in the URIS was the algorithm accuracy not less than the area under the ROC curve (ROC AUC) of 0.81, according to the guidelines for clinical trials of software based on intelligent technologies [14].

Each AI service added a new series of AI-processed CT scans and information in Digital Imaging and Communications in Medicine: Structured Reporting (DICOM SR) format of the study. The additional CT series is based on the original CT series of the current study, with original image supplemented with the segmentation of lung lesions that are caused by COVID-19 according to the AI assumption. The AI developers participating in the experiment were advised to supplement the CT series sent by the AI service with summary information on lung damage and a CT 0–4 score. The DICOM SR data, available to each radiologist in the test group during the formation of the conclusion, contained information about a specific AI service, instructions for using the processing results, and automatically generated report including the severity assessment according to the CT 0–4 scale (Fig. 2).

CT examinations were performed on 48 computed tomographs (Toshiba Aquilion 64, Canon Medical Systems, Japan; HiSpeed GE, USA; Optima CT 660, GE, USA; Somatom Emotion 16, Siemens, Germany; Somatom Sensation 40, Siemens, Germany) according to standard chest scanning protocols recommended by manufacturers.

The comparison was conducted between categories on the CT 0–4 scale from the conclusions of radiologists who had access to AI service results and those who did not have such access.

Main study outcome

AI services were included in the experiment after passing qualitative and quantitative tests on databases prepared by experts of the Scientific and Practical Clinical Center for Diagnostics and Telemedicine Technologies of the Moscow Healthcare Department. Research has been distributed randomly based on the computational capabilities of developers since the inclusion of five AI services. One study could be processed by several AI services. AI results were used only for research purposes, and the radiologist made the final decision on the category on the CT 0–4 scale.

Subgroup analysis

The test group included subgroups since the experiment involved various AI services to diagnose changes in the lungs with COVID-19 according to CT data.

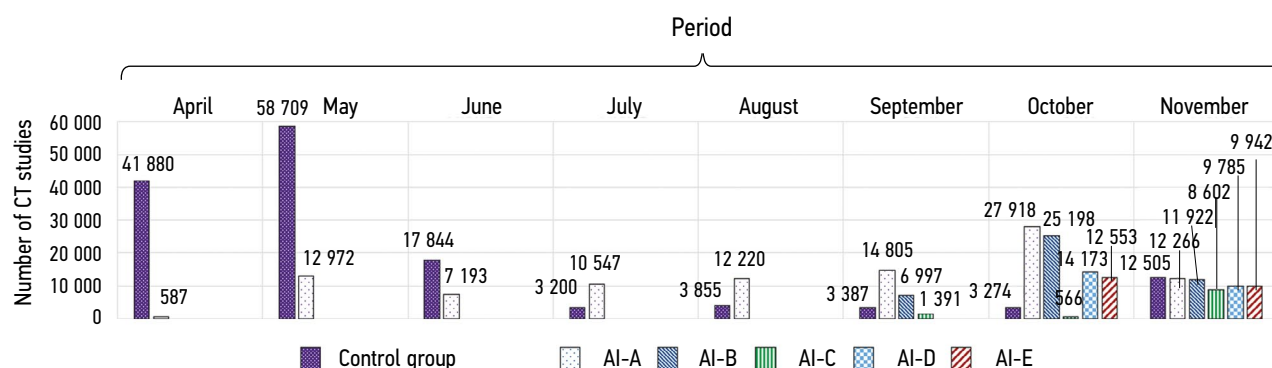


Fig. 1. Chronology of the use of AI services for COVID-19 diagnosis according to computed tomography of the thoracic organs (CT TO).

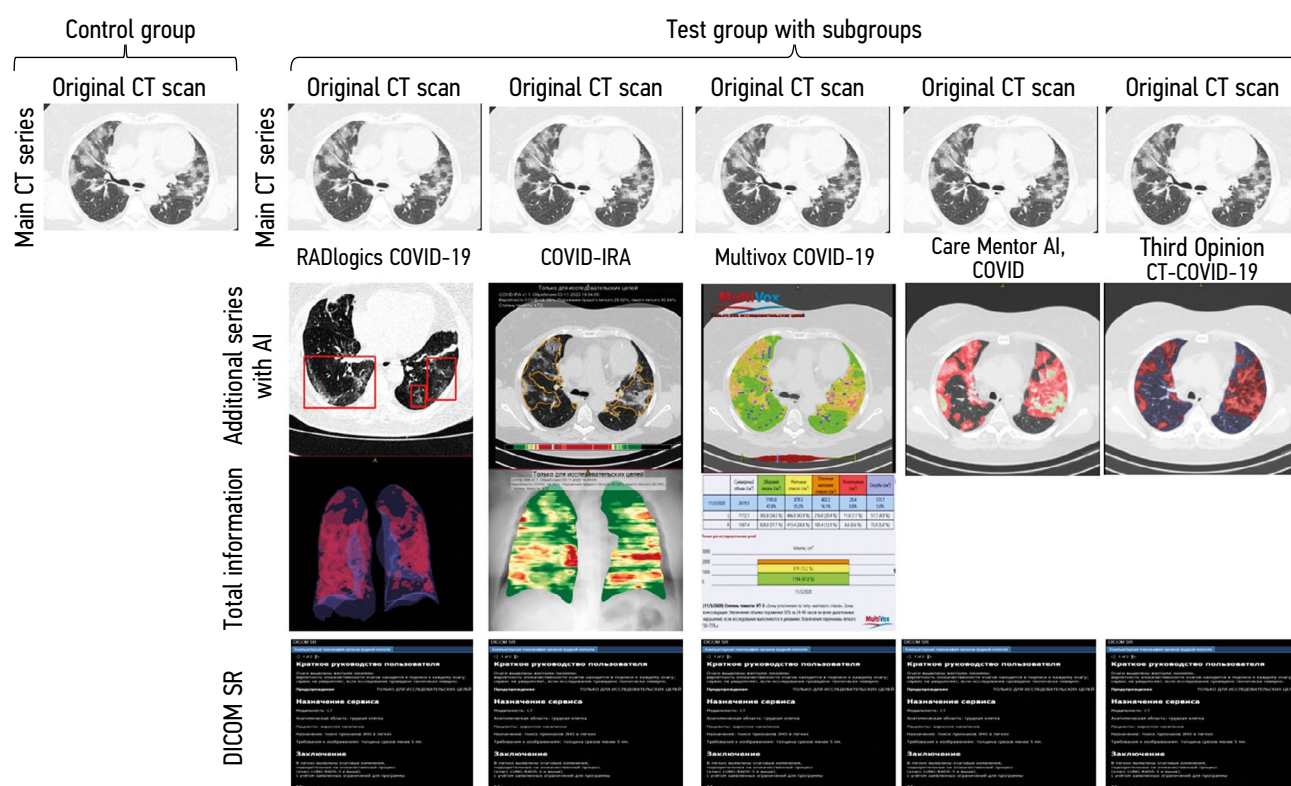


Fig. 2. Examples of original (control group) and additional CT series from various AI services (test group with subgroups) with demonstration of automatic image processing for segmentation of lung lesions in COVID-19, as well as summary information on lung damage and DICOM SR information.

The experiment involved 7 different AI services for diagnosing COVID-19, namely RADlogics COVID-19 (RADLogics, USA), COVID-IRA (IRA labs, Russia), Care Mentor AI, COVID (CareMertor AI, Russia), Third Opinion CT-COVID-19 (Third Opinion, Russia), Multivox COVID19 (Gammamed, Russia), IRYM (Russia), and CVL (CVisionLab, Russia); however, the last two were not included in the test subgroups due to the small number of processed studies over the entire study period. The remaining 5 AI services were randomly coded: AI-A, AI-B, AI-C, AI-D, and AI-E (Fig. 2).

Ethical considerations

Approval of the Independent Ethics Committee of the Moscow Regional Branch of the Russian Society of Roentgenologists and Radiologists was obtained (Protocol No. 2 [1-II-2020] dated February 20, 2020).

Statistical analysis

Principles for calculating the sample size: all those having valid data were included in the statistical analysis. Methods for restoration of missing data were not applied.

Methods for statistical data analysis: descriptive statistics methods were used to present results, indicating the absolute number (n) and proportion (%) of cases in each category. Intergroup comparison of the frequency distribution in different categories between the control group and the test subgroups within each of the 2 periods was performed using the Pearson's chi-squared test (χ^2). The level of statistical

significance was considered as a value of 0.05. Statistical analysis was performed using the Stata 14 software.

RESULTS

Participants of the study

During the entire period, results of primary CT studies of 260,594 patients were analyzed (male/female ratio was 44%/56%, aged from 18 to 100 years, [average age 49.5 years]), performed and interpreted in the period from April 8, 2020 to December 1, 2020 in outpatient medical organizations in Moscow, repurposed for the pandemic period into outpatient CT centers.

Main results of the study

The test group included 115,618 CT studies (44.4% of the total sample), and the control group included 144,976 CT studies (55.6%). The ratio of studies in the control and test groups was similar. Distribution by subgroups in the test group was as follows: 98,953 studies (37.9% of the total sample) were for AI-A, 44,194 (17%) for AI-B, 24,067 (9.2%) for AI-C, 22,679 (8.7%) for AI-D, and 10,645 (4.1%) for AI-E.

For different subgroups with AI services, 2.3%–18.5% fewer CT-0 categories were found (no COVID-19 lesions were detected) compared to the control group, and 4.7%–27.6% fewer CT 3–4 categories was found in different subgroups with AI services than in the control group.

In addition, 40%–60% fewer CT-4 categories were found in different subgroups with AI services than without AI (Fig. 3; $p < 0.0001$).

In November 2020, results of primary CT scans of 41,386 patients were analyzed (male/female ratio was 44%/56%, aged from 18 to 100 years, with the average age of 53.2 years), performed and interpreted in the period from November 1, 2020 to December 1, 2020 in outpatient medical organizations in Moscow, repurposed for the period of the pandemic into outpatient CT centers.

The test group included 28,881 CT scans (69.8% of the total sample in November), and the control group included 12,505 CT scans (30.2%). Distribution by subgroups in the test group was as follows: 12,266 studies (29.6% of the total sample in November) were for AI-A, 11,922 (28.8%) for AI-B, 9,785 (23.6%) for AI-C, 9,942 (24%) for AI-D, and 8,602 (20.8%) for AI-E.

In different subgroups with AI services, 1%–2.6% fewer CT-0 (no COVID-19 lesions were detected), 0.2%–15.7% more CT 3–4, and 25% fewer CT-4 categories were found compared to the control group (Fig. 4; $p = 0.0010$).

For the period of November 2020, statistically significant differences were revealed only for CT-0, whereas no differences were found for the rest of the categories. However, a minimal statistically significant difference was found in CT-0 (18.6% vs. 17.0%). By the general monitoring period, on the contrary, all categories were statistically significantly different between data “without AI” and “total for all AIs”. The critical χ^2 value was 4. For the total period, all CT 0–4 categories made a significant contribution to the differences. The minimum χ^2 value was 26.2 for CT-3 ($p < 0.0001$).

DISCUSSION

Summary of the main result of the study

Study results revealed a change between the CT 0–4 scale categories generated by radiologists in the presence of results from processing by AI services and categories formed without the use of AI services.

Discussion of the main result of the study

In general, prior to the COVID-19 pandemic, AI services were used to reveal radiological symptoms to detect diseases, classify and improve images, reduce radiation exposure, and improve the workflow [15].

Medical research makes AI applications more understandable, safer, more efficient, and more integrated into medical workflows [16]. It should be understood that validation of the AI algorithm should include not only a retrospective analysis of accuracy compared to the golden standard, but also a randomized clinical trial to assess the influence of AI on the decision making by healthcare professionals [17]. For example, in a randomized clinical study HYPE, the effect of the machine learning-based early diagnostics system of intraoperative hypotension was demonstrated, so in the intervention group, the median time of hypotension was 8.0 min versus 32.7 min in the control group ($p < 0.001$), and the number of lethal outcomes in the intervention group was 0 versus 2 in the control group [18]. In another major study using AI to analyze chest X-rays in COVID-19, 20% of surveyed doctors reported that the algorithm influenced clinical decision making [19].

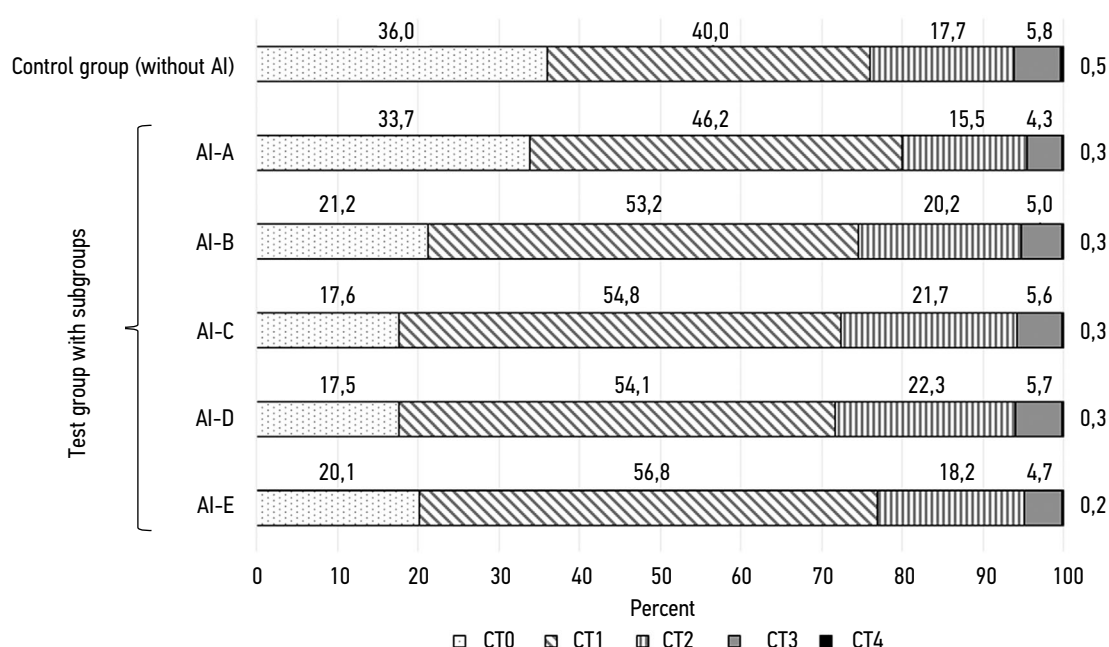


Fig. 3. Results of comparison of primary chest CT scans performed in outpatient CT centers in terms of the severity of CT 0–4 categories between the control group and test subgroups for the entire period (April 8, 2020–December 1, 2020). $n=260\,594$; $p < 0.0001$.

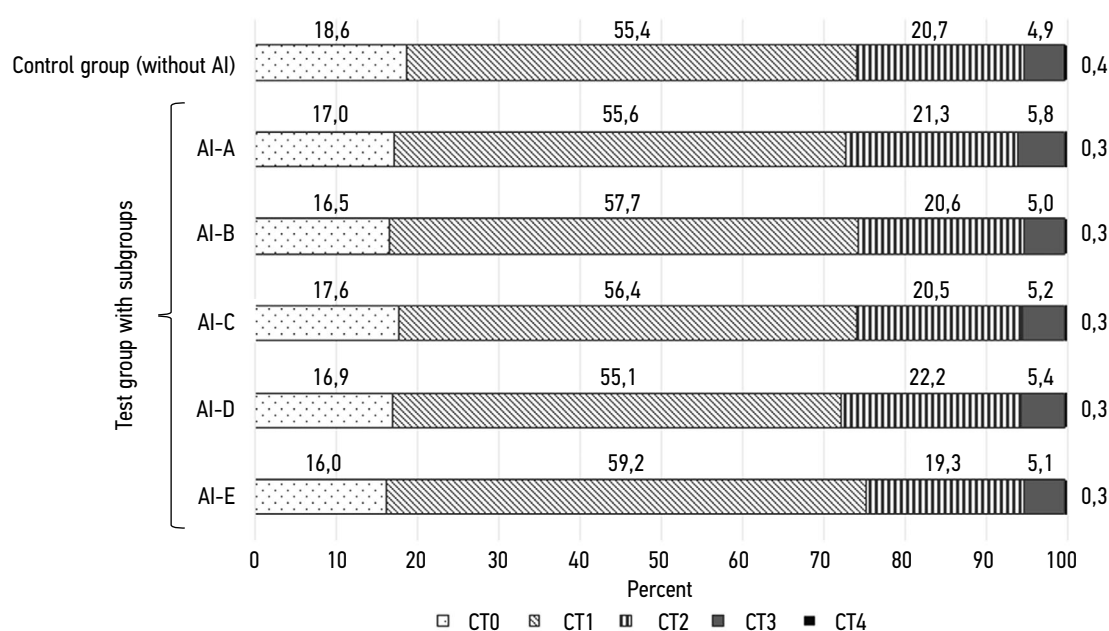


Fig. 4. Results of comparison of primary chest CT scans performed in outpatient CT centers, according to the severity of the CT 0–4 categories between the control group and test subgroups for November 2020.

$n=41\ 386$; $p=0,0010$.

Available literature provides no similar studies assessing the effect of AI results on the final decision making by radiologists when diagnosing lung changes with suspected COVID-19 based on chest CT data.

AI for COVID-19 has two tasks, namely detection and classification. The detection task is determined by the difference between CT-0 and all other categories. The classification task consists of identifying differences between different CT categories (CT1–CT4), i.e., different volumes of damage to the pulmonary parenchyma.

The finding 1 was the difference in the proportion of the CT-0 category between all subgroups. In the case of AI service A, the radiologist was shown only a part of the slices with changes characteristic of COVID-19, while all other AI services had equal number of slices as in the original CT series. Opinions of doctors without AI-A were close to those of the control group.

The finding 2 was the difference in the proportion of CT 3–4 categories between all subgroups. Probably, with a large amount of lung lesions, the empirical visual estimation of the radiologist may exaggerate the extent of the lesion. This overdiagnosis is compensated for when the radiologist monitors the segmentation of lung injuries, performed with the use of AI. This increases the degree of confidence in the automatic calculation of volume and category according to the CT 0–4 scale. Since a lung injury threshold of >50% (CT 3–4 categories) was used as justification for hospitalization, this, combined with clinical and laboratory findings, may reduce the number of hospitalizations.

Based on publications by S.P. Morozov et al. on prediction of lethal outcomes in COVID-19 according to chest CT, when transfer from one CT category to the next, the risk increased by an average of 38% (95% confidence interval

17.1–62.6), and in an additional study among patients with laboratory-verified COVID-19, the risk of lethal outcome with the CT-4 category was 3 times higher than with CT-0 [12, 20]. Our study has demonstrated a fewer CT-4 categories in the test subgroup than in the control one. Previous studies have revealed that CT-4 assessment of the degree of lung damage is associated with lethal outcomes in COVID-19 more than all other categories (CT 0–3) [12, 20]. Thus, the change in the number of patients with CT-4 categories is essential for the formation of calculators of mortality risks for patients with COVID-19.

Due to the pandemic, the simultaneous launch of all AI services was recognized to be limiting the potential benefits of using AI, since all developers would have to be expected to participate in the experiment. Therefore, throughout 2020, unevenness of different AI services in joining the experiment was reported. An analysis was performed for November obtained additional results when the number of CT studies in the control group and each test subgroup was comparable to each other.

According to the authors, the difference revealed between results of periods 1 and 2 is associated with several factors as follows:

- 1) different number of AI services;
- 2) technical factor, as until 2020, the AI services participating in the experiment did not have the opportunity to train their algorithms for assessing and diagnosing lung damage to diagnose COVID-19, therefore, during the experiment, the possibility of changing the AI based software version in order to improve the quality of algorithms and potentially more benefits was recognized justified;
- 3) the human factor, as until 2020, doctors did not use the assessment of the chest CT according to the CT 0–4

scale, which presents a certain difficulty in terms of assessing the volume of multiple lesions in the lungs. It should be noted that doctors could independently improve their skills in assessing the volume of lung lesions in COVID-19, since they evaluated a significant number of CT studies during the pandemic. In addition, doctors could gain experience by checking the markup performed by the AI service, which could lead to an improvement in the skill in more correct assessment of the lesion amount.

In URIS, the radiologist has the opportunity to leave feedback on the work of the AI service in a special feedback field. The study prospect is the comparison of distribution of categories on the CT 0–4 scale among radiologists who have not encountered AI services during the pandemic and use the results of AI services based on their feedback.

Research limitations

Our research has a number of limitations. It did not include patients with positive results of the polymerase chain reaction test for COVID-19 verification, since results of these studies were after the CT scan. The study was not randomized. The extent of agreement of radiologists with the results of AI services was not assessed. In the test group, some of the CT scans were analyzed by several AI services. AI services were not registered as medical devices. Over the course of the pandemic, AI services were changed as the quality of CT processing improved, and this fact was not further evaluated in this study. The adaptation of radiologists to the use of the CT 0–4 scale was not taken into account.

URIS, where doctors formed medical reports, provides a special field for feedback on the work of AI services. However, at the time of the publication formation, results of feedback from doctors were being processed; therefore, it cannot be presented in the current study.

CONCLUSION

Results reveal that the use of AI services for primary chest CT scans in outpatient settings leads to a decrease in

the number of CT-0 and CT 3–4 categories, which can influence the management of patients with COVID-19.

Additional research is required to assess whether reducing the choice of the above categories is appropriate for patient management, and how change in routing further affects recovery and mortality rates.

ADDITIONAL INFORMATION

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. Klyashtorny for statistical analysis.

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 Coronavi-

- rus 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.20200370
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, Gornbolevskiy 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

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.20200370
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-y
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. Синицын 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

AUTHORS' INFO

***Victor A. Gombolevskiy**, Cand. Sci. (Med.);
address: 24/1 Petrovka, Moscow, 127051, Russia;
ORCID: <https://orcid.org/0000-0003-1816-1315>;
eLibrary SPIN: 6810-3279; e-mail: v.gombolevskiy@npcmr.ru

Sergey P. Morozov, Dr. Sci. (Med.), Professor;
ORCID: <https://orcid.org/0000-0001-6545-6170>;
eLibrary SPIN: 8542-1720; e-mail: morozov@npcmr.ru

Anna E. Andreychenko, Cand. Sci. (Phys.-Math.);
ORCID: <https://orcid.org/0000-0001-6359-0763>;
eLibrary SPIN: 6625-4186; e-mail: a.andreychenko@npcmr.ru

Valeria Yu. Chernina, MD;
ORCID: <http://orcid.org/0000-0002-0302-293X>;
eLibrary SPIN: 8896-8051; e-mail: v.chernina@npcmr.ru

Anton V. Vladzimirsky, Dr. Sci. (Med.);
ORCID: <https://orcid.org/0000-0002-2990-7736>;
eLibrary SPIN: 3602-7120; e-mail: a.vladzimirsky@npcmr.ru

Olesya A. Mokienko, Cand. Sci. (Med.);
ORCID: <https://orcid.org/0000-0002-7826-5135>;
eLibrary SPIN: 8088-9921; e-mail: Lesya.md@yandex.ru

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

ОБ АВТОРАХ

Гомболевский Виктор Александрович, к.м.н.;
адрес: Россия, 127051, Москва, ул. Петровка, д. 24/1;
ORCID: <https://orcid.org/0000-0003-1816-1315>;
eLibrary SPIN: 6810-3279; e-mail: v.gombolevskiy@npcmr.ru

Морозов Сергей Павлович, д.м.н., профессор;
ORCID: <https://orcid.org/0000-0001-6545-6170>;
eLibrary SPIN: 8542-1720; e-mail: morozov@npcmr.ru

Андрейченко Анна Евгеньевна, к.ф.-м.н.;
ORCID: <https://orcid.org/0000-0001-6359-0763>;
eLibrary SPIN: 6625-4186; e-mail: a.andreychenko@npcmr.ru

Чернина Валерия Юрьевна;
ORCID: <http://orcid.org/0000-0002-0302-293X>;
eLibrary SPIN: 8896-8051; e-mail: v.chernina@npcmr.ru

Владимирский Антон Вячеславович, д.м.н.;
ORCID: <https://orcid.org/0000-0002-2990-7736>;
eLibrary SPIN: 3602-7120; e-mail: a.vladzimirsky@npcmr.ru

Мокиенко Олеся Александровна, к.м.н.;
ORCID: <https://orcid.org/0000-0002-7826-5135>;
eLibrary SPIN: 8088-9921; e-mail: Lesya.md@yandex.ru