

DOI: <https://doi.org/10.17816/DD321423>

Двойной просмотр результатов маммографии с применением технологий искусственного интеллекта: новая модель организации массовых профилактических исследований

Ю.А. Васильев¹, И.А. Тыров², А.В. Владзимирский¹, К.М. Арзамасов¹,
И.М. Шулькин¹, Д.Д. Кожихина¹, Л.Д. Пестренин¹

¹ Научно-практический клинический центр диагностики и телемедицинских технологий, Москва, Российская Федерация

² Департамент здравоохранения города Москвы, Москва, Российская Федерация

АННОТАЦИЯ

Обоснование. Доступность наборов медицинских данных и технологий разработки программного обеспечения на основе искусственного интеллекта в последние годы привела к увеличению количества решений для медицинской диагностики и маммографии в частности. Это программное обеспечение, зарегистрированное как медицинское изделие, может быть использовано для описания цифровых маммографий, что позволит в значительной мере сэкономить временные, материальные и кадровые ресурсы в здравоохранении при гарантированном сохранении качества профилактических исследований молочных желёз.

Цель — обосновать возможность и эффективность применения программного обеспечения на основе технологий искусственного интеллекта для первой интерпретации цифровых маммограмм при сохранении практики второго описания рентгеновских изображений врачом-рентгенологом.

Материалы и методы. Набор данных из 100 цифровых маммографических исследований, из них 50 — «Отсутствие целевой патологии», 50 — «Присутствие целевой патологии» (с признаками злокачественных новообразований), был обработан программным обеспечением на основе технологий искусственного интеллекта, зарегистрированным в Российской Федерации как медицинское изделие. Выполнен ROC-анализ. Ограничения исследования: значения метрик диагностической точности получены для версий программного обеспечения на основе технологий искусственного интеллекта, актуальных на конец 2022 года.

Результаты. При настройке на 80,0% чувствительность специфичность искусственного интеллекта составила 90,0% (95% ДИ 81,7–98,3), точность — 85,0% (95% ДИ 78,0–92,0). При настройке на 100% специфичность искусственный интеллект показал чувствительность 56,0% (95% ДИ 42,2–69,8), точность — 78,0% (95% ДИ 69,9–86,1). При настройке на 100% чувствительность специфичность искусственного интеллекта составила 54,0% (95% ДИ 40,2–67,8), точность — 77,0% (95% ДИ 68,8–85,2). Предложены два подхода, предусматривающие автономную первую интерпретацию цифровой маммографии посредством искусственного интеллекта. Первый подход заключается в оценке рентгеновского изображения с помощью искусственного интеллекта с более высокой чувствительностью, чем у двойного описания маммографии врачами-рентгенологами, при сопоставимом уровне специфичности. Второй подход подразумевает, что программное обеспечение на основе технологий искусственного интеллекта будет определять категорию маммографии («Отсутствие целевой патологии» или «Присутствие целевой патологии») с указанием степени своей «уверенности» в полученном результате в зависимости от «коридора», в который попадает предсказанное значение.

Заключение. Оба предложенных сценария использования программного обеспечения на основе технологий искусственного интеллекта с целью автономного первого описания цифровых маммограмм способны обеспечить качество диагностики, не уступающее двойному описанию снимков врачами-рентгенологами и даже превышающее его. Экономическая выгода от практической реализации данного подхода в масштабах страны может составлять от 0,6 до 5,5 млрд рублей ежегодно.

Ключевые слова: профилактические исследования; маммография; искусственный интеллект; диагностическая точность.

Как цитировать

Васильев Ю.А., Тыров И.А., Владзимирский А.В., Арзамасов К.М., Шулькин И.М., Кожихина Д.Д., Пестренин Л.Д. Двойной просмотр результатов маммографии с применением технологий искусственного интеллекта: новая модель организации массовых профилактических исследований // *Digital Diagnostics*. 2023. Т. 4, № 2. С. 93–104. DOI: <https://doi.org/10.17816/DD321423>

Рукопись получена: 17.03.2023

Рукопись одобрена: 25.04.2023

Опубликована: 25.05.2023

DOI: <https://doi.org/10.17816/DD321423>

Double-reading mammograms using artificial intelligence technologies: A new model of mass preventive examination organization

Yuriy A. Vasilev¹, Ilya A. Tyrov², Anton V. Vladzimirskyy¹, Kirill M. Arzamasov¹, Igor M. Shulkin¹, Daria D. Kozhikhina¹, Lev D. Pestrenin¹

¹ Moscow Center for Diagnostics and Telemedicine, Moscow, Russian Federation

² Moscow Health Care Department, Moscow, Russian Federation

ABSTRACT

BACKGROUND: In recent years, the availability of medical datasets and technologies for software development based on artificial intelligence technology has resulted in a growth in the number of solutions for medical diagnostics, particularly mammography. Registered as a medical device, this program can interpret digital mammography, significantly saving time, material, and human resources in healthcare while ensuring the quality of mammary gland preventive studies.

AIM: This study aims to justify the possibility and effectiveness of artificial intelligence-based software for the first interpretation of digital mammograms while maintaining the practice of a radiologist's second description of X-ray images.

MATERIALS AND METHODS: A dataset of 100 digital mammography studies (50 — “absence of target pathology” and 50 — “presence of target pathology,” with signs of malignant neoplasms) was processed by software based on artificial intelligence technology that was registered as a medical device in the Russian Federation. Receiver operating characteristic analysis was performed. Limitations of the study include the values of diagnostic accuracy metrics obtained for software based on artificial intelligence technology versions, relevant at the end of 2022.

RESULTS: When set to 80.0% sensitivity, artificial intelligence specificity was 90.0% (95% CI, 81.7–98.3), and accuracy was 85.0% (95% CI, 78.0–92.0). When set to 100% specificity, artificial intelligence demonstrated 56.0% sensitivity (95% CI, 42.2–69.8) and 78.0% accuracy (95% CI, 69.9–86.1). When the sensitivity was set to 100%, the artificial intelligence specificity was 54.0% (95% CI, 40.2–67.8), and the accuracy was 77.0% (95% CI, 68.8–85.2). Two approaches have been proposed, providing an autonomous first interpretation of digital mammography using artificial intelligence. The first approach is to evaluate the X-ray image using artificial intelligence with a higher sensitivity than that of the double-reading mammogram by radiologists, with a comparable level of specificity. The second approach implies that artificial intelligence-based software will determine the mammogram category (“absence of target pathology” or “presence of target pathology”), indicating the degree of “confidence” in the obtained result, depending on the corridor into which the predicted value falls.

CONCLUSIONS: Both proposed approaches for using artificial intelligence-based software for the autonomous first interpretation of digital mammograms can provide diagnostic quality comparable to, if not superior to, double-image reading by radiologists. The economic benefit from the practical implementation of this approach nationwide can range from 0.6 to 5.5 billion rubles annually.

Keywords: artificial intelligence; diagnostic accuracy; mammography; preventive medicine.

To cite this article

Vasilev YuA, Tyrov IA, Vladzimirskyy AV, Arzamasov KM, Shulkin IM, Kozhikhina DD, Pestrenin LD. Double-reading mammograms using artificial intelligence technologies: A new model of mass preventive examination organization. *Digital Diagnostics*. 2023;4(2):93–104. DOI: <https://doi.org/10.17816/DD321423>

Received: 17.03.2023

Accepted: 25.04.2023

Published: 25.05.2023

DOI: <https://doi.org/10.17816/DD321423>

使用人工智能对乳腺X线摄影结果进行两次查看：一种组织大规模预防性研究的新模式

Yuriy A. Vasilev¹, Ilya A. Tyrov², Anton V. Vladzimirskyy¹, Kirill M. Arzamasov¹, Igor M. Shulkin¹, Daria D. Kozhikhina¹, Lev D. Pestrenin¹¹ Moscow Center for Diagnostics and Telemedicine, Moscow, Russian Federation² Moscow Health Care Department, Moscow, Russian Federation

简评

论证。近年来，医疗数据集和人工智能软件技术开发的可达性，使得医疗诊断，特别是乳腺X线摄影的解决方案激增。这种登记为医疗设备的软件可被用于描述数乳腺X线摄影，这允许提供医疗服务时在很大程度上节省时间、物质和人力的资源，同时确保乳房预防性检查的质量。

该研究的目的是证明使用基于人工智能技术的软件对数字乳房X光照片进行第一次解读的可用性和有效性，同时保持放射科医生对X射线图像进行第二次描述的做法。

材料和方法。我们用在俄罗斯联邦登记为医疗设备的基于人工智能技术的软件处理了100张数字乳房X光照片的数据集，其中50张为“无目标病变”，50张为“存在目标病变”（有恶性肿瘤症状）。进行了ROC分析。研究局限性：诊断准确性度量值是基于人工智能技术的软件版本获得的，是在2022年底有效的。

结果。当设置为80.0%的灵敏度时，人工智能显示出90.0%的特异度（95% CI 81.7–98.3）和85.0%的准确性（95% CI 78.0–92.0）。当设置为100%的特异度时，人工智能显示出56.0%的灵敏度（95% CI 42.2–69.8）和78.0%的准确性（95% CI 69.9–86.1）。当设置为100%灵敏度时，人工智能的特异度为54.0%（95% CI 40.2–67.8），准确性为77.0%（95% CI 68.8–85.2）。提出了两种方法，涉及通过人工智能对数字乳腺X线摄影进行的第一次自主解读。第一种方法是利用人工智能评估X射线图像，其灵敏度高于放射科医生进行的双重乳腺X线摄影描述，特异度水平相当。第二种方法是基于人工智能技术的软件将对乳腺X线摄影进行分类（“无目标病变”或“存在目标病变”），表明其对结果的“信心”程度，取决于预测值所处的“走廊”。

结论。使用基于人工智能技术的软件对数字乳房X光照片进行第一次自主描述的两方案都能提供与放射科医生对图像进行的双重描述相同甚至更高的诊断质量。在全国范围内在实践中实现这种方法的经济效益可能是每年6亿至55亿卢布。

关键词：预防性研究；乳腺X线摄影；人工智能；诊断准确性。

引用本文：

Vasilev YuA, Tyrov IA, Vladzimirskyy AV, Arzamasov KM, Shulkin IM, Kozhikhina DD, Pestrenin LD. 使用人工智能对乳腺X线摄影结果进行两次查看：一种组织大规模预防性研究的新模式. *Digital Diagnostics*. 2023;4(2):93–104. DOI: <https://doi.org/10.17816/DD321423>

收到: 17.03.2023

接受: 25.04.2023

发布日期: 25.05.2023

论证

乳腺恶性肿瘤是一个重要的医疗、社会经济和人口问题，作为女性人口中主要肿瘤病变及死亡率最高的疾病。从2011年到2019年，发病率稳步上升（分别从每10万人口45.24到53.34）；2020年急剧下降到每10万人口47.39，2021年又恢复上升[1]。这种动态是因为在COVID-19大流行期间大规模预防性检查项目被暂停。一方面，这种情况是完全合理的；另一方面，它显示医疗保健系统的脆弱性。被迫的资源重新分配对“社会意义”疾病的筛查产生了负面影响。但是，在没有大流行病的情况下，乳腺恶性肿瘤检测的及时性仍然不能令人满意。忽视率极高：超过27.0%的恶性肿瘤在晚期（III-IV期）被发现。有一个积极的趋势引人注目——乳腺恶性肿瘤的累积指数稳步上升：2011年适当指标是9.5，2021年已经是11.9。这表明，乳腺恶性肿瘤治疗方法的质量和效果有系统地提高[2]。

因此，需要发展大规模的预防性研究，既要提高人口覆盖率和参与率、生产力、质量、成本效益，又要建立持续性和不间断的可及性。这些任务的实施与治疗方法的明显进步相结合，将使克服乳腺恶性肿瘤问题获得完全新的成果。

目前乳腺X线摄影是最常见的恶性肿瘤筛查方式。

现行规范性文件规定，预防性乳腺X线摄影的结果必须经过两次审查。这就是说每个受检患者的图像必须由两位放射科医生独立解释和描述。这种方法的合理性已被国内外学者反复证实。当两次审查图像时发现病理变化的综合频率更高。与两次审查相比，一次审查会降低所有BI-RADS (Breast Imaging Reporting and Data System, 乳腺影像报告和数据库系统) 类别的灵敏度。此外，放弃两次审查会给受检妇女带来各种负面影响[3, 4]。同时，两次审查也有一定的负面作用（资源密集度、质量、不完善的筹资系统）。

资源密集度。在初级卫生保健环境中，需要两名放射科医生来描述每项预防性检查，任何筛查检查的绝大部分都属于“无目标病变”类别。由于浮报和其他操纵手法，存在着隐瞒人才断层的风险，这最终对女性人口的健康产生了负面影响。同时，预防性乳腺X线摄影数量的实际需求和设备存量的速度增加从根本上说高于增加工作人员的能力。

质量。解释乳腺X线摄影的结果需要特殊技能，实际上是现代放射学的一个亚专业。这只会加剧人才断层的问题：“机械地”增加放射科医生的数量并不能解决按质和及时发现乳腺恶性肿瘤的问题。

不完善的筹资系统。使用两部收费，即连续支付两项服务的费用，只适用于有单独的乳腺X线摄影描述收费的情况；这样一来，一项服务包括乳腺X线摄影的进行和描述，而另一项只包括其描述。在相反的情况下，存在不适当的筹资（同

时包括乳腺X线摄影进行和描述的两项服务被支付）。一个更常见的情况是医疗机构资金不足：进行和描述的服务被支付一次，而第二次描述则没有被支付。

这些问题有可能通过自动化来解决，即使用人工智能(AI)技术来解释乳腺X线摄影的结果。

在科学文献中，有证据表明所涉及的技术解决方案的质量很高。事实上，在检测某些放射学特征方面，一些基于人工智能技术的描述诊断准确度与普通放射科医生相同。因此，自动分析的合并灵敏度、特异性和特征曲线下的面积为75.4；90.6和0.89%，而放射科医生的为73.0；88.6和0.85%（没有统计学意义的差异）[5]。根据荟萃分析的数据[6]，对乳腺X线摄影进行自动分析的合并灵敏度、特异性和特征曲线下的面积分别为91.4；91.6和94.5%。神经网络模型和放射科医生检测到的病理检查所见的定位的IoU-准确度（交并比，intersection-over-union）相似系数分别为0.86和0.96[7]。然而，我们认为，在乳腺恶性肿瘤筛查的背景下，人工智能技术不应仅仅作为医生的离散决策支持系统，而应作为自动程序来引入。

该研究的目的是证实一个利用人工智能技术组织大规模预防性乳腺X线摄影的模式。

材料和方法

该研究是莫斯科政府从2020年开始支持的《使用创新计算机视觉技术进行医学图像分析并进一步应用于莫斯科市医疗系统的实验》（以下简称莫斯科实验）(mosmed.ai)的一部分。基于人工智能技术的软件所达到的准确度(AUC, 曲线下的面积>0.81)，是在莫斯科实验第一年内对61 497张乳房X光照片的分析基础上调定的，这使得该研究成为可能[8]。

研究设计

混合型（定量部分为回顾性诊断研究，定性部分为分析研究）。

人工智能技术

为了在临床实践中使用，基于人工智能技术的软件的记录必须显示为医疗器械。在俄罗斯联邦，Medical Screening Systems LLC (RZN 2021/14449) 和Third Opinion Platform LLC (RZN 2022/16534) 基于人工智能技术的软件产品被用于对乳腺X线摄影的结果进行自动分析，具有医疗器械的地位。

研究中使用的医疗器械是上述公司之一的基于人工智能技术的软件产品。作为莫斯科实验的一部分，基于人工智能技术的软件被整合到莫斯科市统一医疗信息分析系统的统一放射学信息系统(ERIS EMIAS)中。基于人工智能技术的软件接收了DICOM格式的数字乳腺X线摄影，对其进行了分析，并将文本描述(DICOM SR)、带有病理

区域标记的图像 (DICOM SC) 以及检查中存在病理的概率作为回复。基于人工智能技术的软件进行的分析结果与原始检查结果都是在ERIS EMIAS中提供的。研究中使用了整个研究中的病理概率值 (存在恶性过程的概率); 在本研究中对BI-RADS的正确性以及病理检查所见的定位准确性没有进行评估。

数据集

我们的研究共包括100张作为筛查研究一部分的数字乳房X光照片, 其中50张是有病理的。接受检查的患者的平均年龄为 63 ± 6 岁。

根据两位专家 (具有5年以上乳腺X线摄影经验的放射科医生) 的共识, 这些研究被分为“无目标病变”或“存在目标病变”。对于“存在目标病变”组来说, 组织学验证是主要的纳入标准。对于“无目标病变”组来说, 放射科医生的不同意见是排除标准。两组的排除标准都是患者年龄小于18岁和研究质量差 (PGMI分级为I)。质量差的研究是在准备数据集时审查员进行标记阶段中确定的。

与BI-RADS3、4和5级相对应的在乳房X光照片上的病理征被认为是目标病理 (“存在目标病变”)。研究中纳入“无目标病理”组的乳房X光照片是没有乳腺恶性肿瘤的, 即BI-RADS分级为1或2级。

根据ACR (美国放射学会, American College of Radiology) 分类, 研究分配为: “存在目标病变”组中: A-26, B-16, C-5, D-3; “无目标病变”组中: A-15, B-24, C-11, D-0。

该数据集是由在FUJIFILM公司 (日本) 乳腺X射线机上进行的研究组成的。以下医疗机构参与了数据集的准备: 城市综合医院第22号, 第1分院; 临床诊断中心第1号; 城市综合医院第8号; 城市综合医院第36号; 城市综合医院第22号; 城市综合医院第209号; 诊断中心第3号, 第4分院; 咨询和诊断综合医院第121号; 临床诊断中心第4号; 莫斯科卫生局M. P. Konchalovsky城市临床医院, 综合医院第3部。

伦理审查

本研究建立在《使用创新计算机视觉技术进行医学图像分析并进一步应用于莫斯科市医疗系统的实验》结果的基础上, 该实验由伦理委员会批准 (自2020年2月20日的协议№2 NEC MRO RORR的摘抄), 也在ClinicalTrials (NCT04489992) 登记注册。

统计分析

为了处理获得的数据, 通过专门开发的网络工具 (<https://roc-analysis.mosmed.ai/>) [9, 10], 采

用了线绘和分析特征曲线 (ROC分析) 的方法。真值由研究的二元划分表示 (0为无病变, 1为存在目标病变)。结果是由基于人工智能技术的软件确定的病理存在的概率。将获得的数以CSV格式上传到电子表格, 然后用网络工具线绘特征曲线。以交互方式确定与最左边的100%灵敏度的点和最右边的100%特异性的点对应的阈值。然后对确定的阈值进行了其他诊断准确度指标的分析。典型四格表的术语被用于分析。真阳性被认为是将乳房X光照片正确地分配到“存在目标病理”组; 真阴性被认为是将乳房X光照片正确地分配到“无目标病理”组。假阳性被认为是将乳房X光照片从“正常”组错误地分配到“存在目标病变”组, 而假阴性被认为是将乳房X光照片从“存在目标病变”组错误地分配到“无目标病变”组。

结果中所提出的统计指标都是通过有返回的随机抽样 (所谓的bootstrapping), 经过1000次迭代计算出95%的置信区间 (CI)。

结果

我们提出以下模式。预防性乳腺X线检查是根据现行规范性文件 and 临床建议由主治医师而指定的。X线实验员直接进行检查。检查结果以数字方式输入医学图像档案中, 该档案是医疗机构的医疗信息系统和/或俄罗斯联邦主体卫生领域信息系统的一部分。第一次查看是由软件 (基于人工智能技术的医疗器械, 已被批准按既定顺序流通) 进行的。进行第一次查看之后, 在信息系统中形成以自动模式获得的电子医疗记录¹, 并包含 (a) 带有有病理变化区域的图形标记和/或温度图的独立系列图像 (如果有病理变化的话); (b) 结构化报告、简要用户指南、体检结论、详细说明和存在病理变化的概率。第二次查看是由放射科医生进行的。进行第二次查看之后, 在信息系统中形成一份电子医疗文件, 即描述协议和体检结论。

对基于人工智能技术的软件提供所需诊断准确度水平的可行性进行了研究。

图1给出描写研究中的基于人工智能技术的软件工作的特征曲线。

根据选定的阈值将乳房X光照片分为几组 (真阳性、真阴性、假阴性、假阳性), 表1给出其分布情况。

当定出与100.0%特异性相对应的阈值 (阈值为0.93) 时, 人工智能正确识别了“存在目标病理”组的乳房X光照片, 即没有假阳性。在50张属于“存在目标病理”组的乳房X光照片中, 有28张 (56.0%) 被人工智能在指定的阈值设置下正确分配到。对于这些基于人工智能技术的软件设置来

¹ 因使用获准流通的医疗器械而以自动模式获得的医疗记录不需要医务人员的电子签名 (根据俄罗斯联邦卫生部于2020年9月7日第947n号命令《关于批准保健领域以电子文件形式的医疗记录的文件管理系统组织程序》)。

说, 诊断准确度指标为: 灵敏度为56.0% (95% CI 42.2–69.8), 特异性为100.0% (95% CI 100.0–100.0), 以及准确度为78.0% (95% CI 69.9–86.1)。

当定出与100.0%灵敏度相对应的阈值 (阈值为0.25) 时, 没有假阴性并真阴性数量为27例 (占“无目标病理”组所有测试的54.0%)。对于这些基于人工智能技术的软件设置来说, 诊断准确度指标为: 灵敏度为100.0% (95% CI 100.0–100.0), 特异性为54.0% (95% CI 40.2–67.8), 以及准确度为77.0% (95% CI 68.8–85.2)。

当使用约登指数最大值法定出阈值 (阈值为0.82) 时, 灵敏度为80.0%, 真阴性为45

例 (占“无目标病理”组所有测试的90.0%), 真阳性为40例 (占“存在目标病理”组所有测试的80.0%)。对于这些基于人工智能技术的软件设置来说, 诊断准确度指标为: 灵敏度为80.0% (95% CI 68.9–91.1), 特异性为90.0% (95% CI 81.7–98.3), 准确度为85.0% (95% CI 78.0–92.0)。

讨论

该研究的主要结果摘要

当实现《使用人工智能描述和解释乳腺X线检查数据》服务时使用了基于人工智能技术的软件, 定出不同的阈值允许指定灵敏度、特异性和

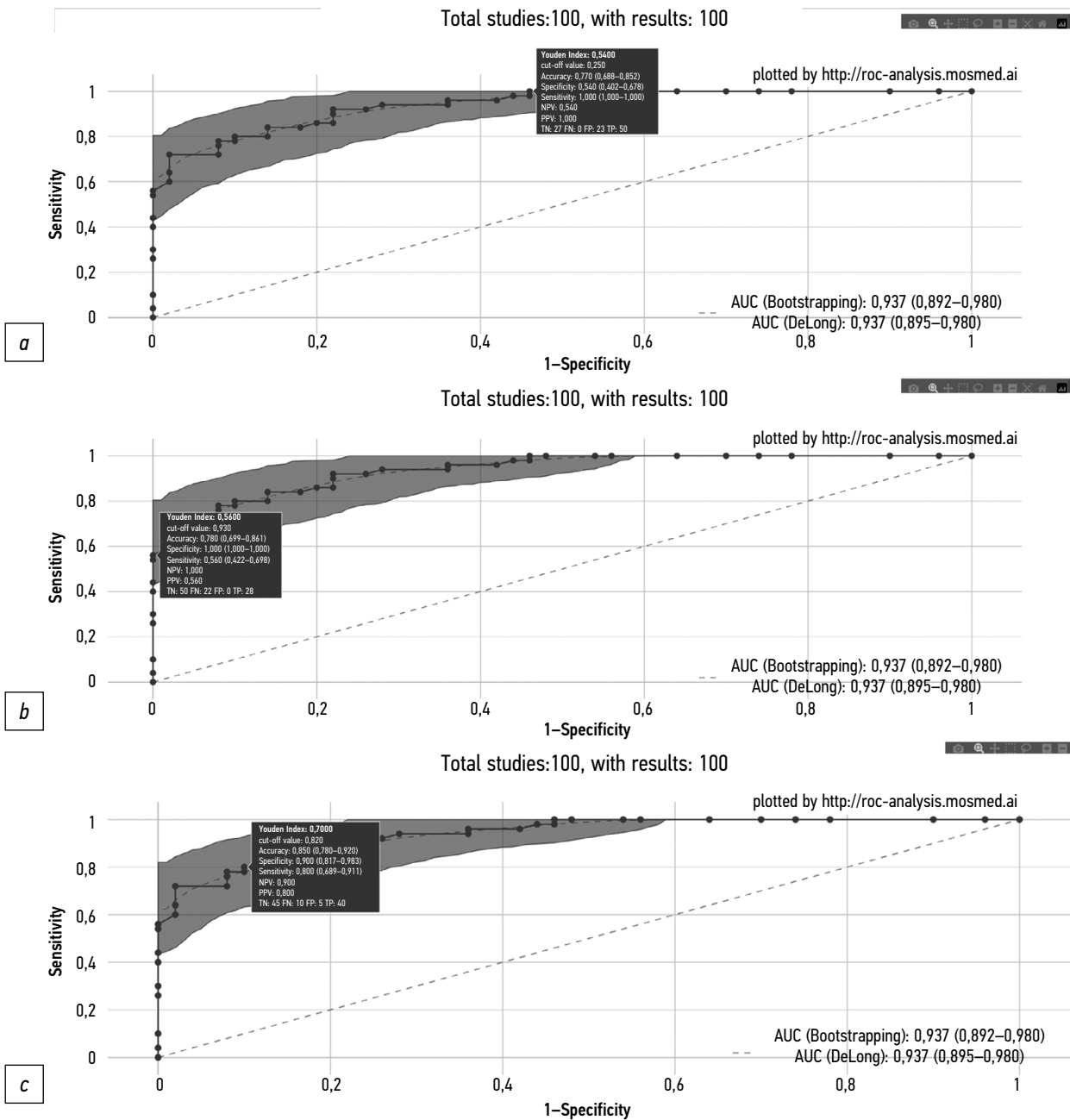


图1. 基于人工智能技术软件的特征曲线。底纹表示95%的置信区间。分别显示了对应于100.0%灵敏度 (a)、100.0% 特异性 (b) 和80.0%灵敏度 (c) 的实验值。对于每一个实验点, 在相应的阈值处的诊断准确度指标的数值都显示在一个矩形中。

表1. 不同阈值的四格表

组	研究的数量	结果	100%的灵敏度	100%的特异性	平衡的灵敏度和特异性
“存在目标病理”	50	真阳性	50	28	40
		假阳性	23	0	5
“无目标病理学”	50	真阴性	27	50	45
		假阴性	0	22	10

注: *真阴性的百分比是相对于“无目标病理”组的规模计算的; **真阳性的百分比是相对于“存在目标病理”组的规模计算的。

准确度, 这些指标与放射科医生对乳房X光照片进行两次查看后获得的数值一致或超过该数值。

对该研究主要结果的讨论

在俄罗斯联邦, 每年大约有820万个预防性乳腺X线检查², 对其进行双重评估需要大量的时间、人力和财政资源。使用基于人工智能技术的软件对乳房X光照片进行第一次查看允许减少这些成本, 同时保持甚至提高诊断质量。建议采用两种调整基于人工智能技术软件的方法来对乳房X光照片进行第一次查看。

第一种方法是使用基于人工智能技术的软件, 其灵敏度和特异性的数值是平衡的。在我们的研究中, 基于人工智能技术软件的灵敏度为80.0%, 超过了综述性出版物中指出的放射科医生对乳房X光照片进行两次查看的灵敏度(72.0–73.0%) [5, 6]。然而, 我们研究中发现的人工智能的特异性(90.0%)与两位放射科医生工作的特异性(88.0–98.0%)相当[5, 6]。我们也预计, 与单独一名放射科医生的描述相比, 与一名放射科医生的评估相结合的基于人工智能技术的软件会有更高的乳房X光照片解释的总体准确度, 这也是在一些科学出版物中证实的[11–13]。在这种方法中, 基于人工智能技术软件的结果将是一份电子医疗记录, 其中包含对图像类别(“无目标病变”或“存在目标病变”)的意见。

第二种方法意味着基于人工智能技术的软件将确定图像类别(“无目标病变”或“存在目标病变”), 表明其对结果的“信心”程度。图2给出该方法的一般概念。

如《结果》部分所述, 在线绘基于人工智能技术软件的ROC曲线时, 我们确定了100%灵敏度和100%特异性下预测值的阈值(cut-off value) (分别为0.25和0.93)。在这些数据的基础上, 建议将预测值分配到三个“走廊”, 这三个“走廊”适当不同的分类结果和基于人工智能技术的软件对“走廊”的不同“信心”程度:

1) “绿色走廊”: 预测值从0到0.25不等, 对应于“无目标病理”类别, 有100%的把握;

2) “红色走廊”: 预测值从0.93到1.0不等,

对应于“存在目标病理”类别, 有100%的把握;

3) “黄色走廊”: 预测值从0.25到0.93不等(含0.93), 对应于“无目标病变”或“存在目标病变”的类别, 但正确分类的概率小于100%。

建议在使用基于人工智能技术的软件进行的乳腺X线检查描述中加入预测值和其所处的“走廊”颜色。有了这些信息, 描述基于人工智能技术的软件图像的放射科医生将知道对结果的信心有多大。一方面, 这有助于医生在描述“黄色走廊”的乳房X光照片时保持高度警惕, 另一方面, 从长远来看, 这种方法可能会提升放射科医生对基于人工智能技术的软件结果的信心。这是因为即使目前人工智能的准确度很高, 人工智能也不能明确地对所分析的图像进行100%的正确分类。

第二种方法的优势不仅在于人工智能能够以100%的“信心”程度将一些乳房X光照片分类为“无目标病变”或“存在目标病变”(预测值的“绿色和红色走廊”), 而且在于能够改变阈值, 以平衡基于人工智能技术的软件对属于“黄色走廊”的乳房X光照片进行的分析灵敏度和特异性。根据临床目标, 可以指定更高的灵敏度, 以假阴性的数量最少时提供更好的病理检测, 或者指定更高的特异性, 以减少假阳性(表2) [5, 6]。

本研究的结果显示使用基于人工智能技术的软件对乳房X光照片进行第一次查看的可行性; 然而, 该软件还需要进一步优化, 以更好地识别“无目标病变”和“存在目标病变”。还值得注意的是, 为了进行本研究, 最初只选择了高质量的乳房X光照片。然而, 基于人工智能技术的现代软件具有控制乳房X光照片质量的自主功能。在实际实施中, 除了对乳腺X线检查结果进行临床评估外, 基于人工智能技术的软件还可以对质量进行技术评估[8]。

使用人工智能技术对乳腺X线检查结果进行两次查看的经济论证

该研究中我们分析了俄罗斯联邦各地区通过的2023年强制医疗保险领土方案下提供的医疗服务的支付费率协议。

² Tyurin I. E. 俄罗斯卫生部辐射和仪器诊断首席外部专家I. E. Tyurin的2020年报告[电子资源]。访问方式: https://static-0.minzdrav.gov.ru/system/attachments/attaches/000/056/620/original/О_т_ч_е_т_з_а_2_0_2_0_г_о_д_Т_ю_р_и_н_.pdf?1624967722。

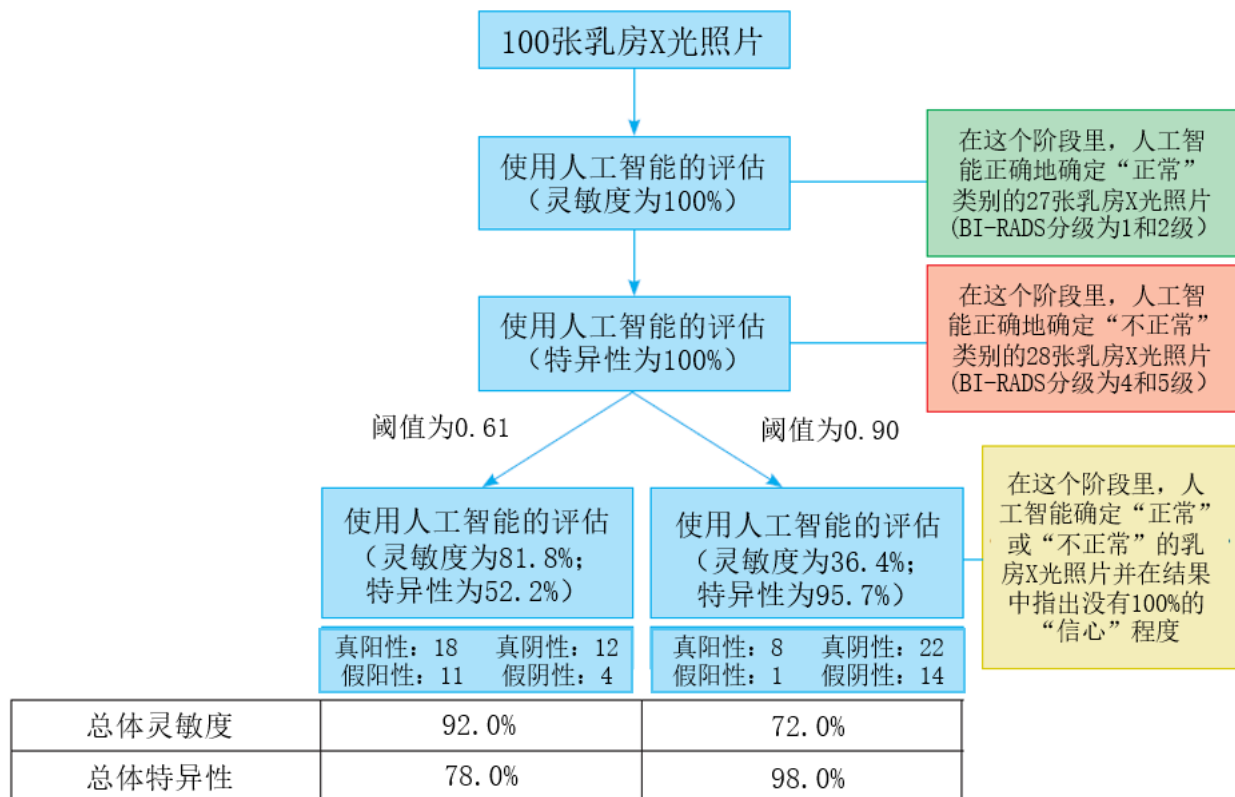


图2. 使用人工智能的乳腺X线检查首次描述方法的概念，该方法基于二元图像划分，表明基于人工智能技术的软件对所获结果的“信心”程度。

在俄罗斯联邦的19个主体（占俄罗斯联邦所有主体的22.4%）中，单独确定了在临床检查框架内进行的《A06.20.004 乳腺X线摄影》医疗服务的费用。19个主体中的4个主体也确定了《A06.30.002 X射线图像的描述和解释（在重新描述乳房X光照片的情况下）》医疗服务的费用。在俄罗斯联邦的所有其他主体中，成人健康检查的第一阶段被算作一项综合服务，而没有明确包括在内的乳腺X线检查的费用。

乳腺X线检查描述的医疗服务费用从114.97到1034.93卢布不等。截止到2023年3月1日，《使

用人工智能描述和解释乳腺X线检查数据》医疗服务只是在莫斯科实验的框架内应用的[8]。根据莫斯科市强制医疗保险领土方案下提供的医疗服务的支付费率协议，该医疗服务的费用为239.00卢布³。

该研究中我们分析了两种方法来确定俄罗斯联邦境内进行筛查乳腺X线摄影的必要资金。第一种方法是根据2023年莫斯科的医疗服务成本进行计算。放射科医生进行一次乳腺X线检查描述的费用是178.00卢布。因此，放射科医生对每张乳房X光照片进行的两次查看值356.00卢布。反过

表2. 不同乳房X光照片分析方法的灵敏度和特异性

预防性乳腺X线摄影结果的分析	灵敏度, %	特异性, %
由医生进行的两次查看*	72.0 - 73.0	88.0 - 98.0
第一种使用人工智能对乳房X光照片进行第一次查看的初读的方法（二元划分）	80.0	90.0
第二种使用人工智能对乳房X光照片进行第一次查看的方法（具有“置信度”的二元划分），阈值为0.61	92.0	78.0
第二种使用人工智能对乳房X光照片进行第一次查看的方法（具有“置信度”的二元划分），阈值为0.90	72.0	98.0

注: *基于文献[5, 6]。

³ 2023年莫斯科市强制医疗保险领土方案下提供的医疗服务的支付费率协议（2022年12月30日在莫斯科签订）。访问方式：<https://www.consultant.ru/law/review/208575247.html>。

来, 如上所述, 人工智能和放射科医生描述乳房X光照片的费用为239.00卢布。因此, 在俄罗斯每年平均有820万张乳房X光照片的情况下, 放射科医生进行两次描述的费用将是29亿卢布, 而人工智能和放射科医生进行描述的费用将是19亿卢布。使用基于人工智能技术的软件, 每年可能节省10亿卢布。第二种方法是根据2023年俄罗斯联邦各主体的医疗服务成本进行计算。同时还考虑了人工智能和放射科医生在莫斯科的乳腺X线检查解释与放射科医生的两次解释相比所节省的成本百分比, 即

$$\frac{178.00 \times 2 - 239.00}{178.00 \times 2} \times 100\% = 32.8\%.$$

在俄罗斯联邦各地区放射科医生描述乳房X光照片的医疗服务费用为114.97至1034.93卢布, 而放射科医生两次描述乳房X光照片的费用为229.94至2069.86卢布。假设俄罗斯联邦各地区的人工智能和放射科医生对乳房X光照片的解释与莫斯科相同, 并比放射科医生的两次描述便宜32.8%, 我们得出人工智能和放射科医生对乳房X光照片的描述费用在154.51至1390.94卢布之间。因此, 在俄罗斯每年平均有820万张乳房X光照片的情况下, 放射科医生进行两次描述的成本将是18至169亿卢布, 而人工智能和放射科医生的描述成本将是12至114亿卢布。在全国范围内使用基于人工智能技术的软件, 每年可能会节省6亿至55亿卢布。

该研究的局限性

诊断准确度指标的数值是针对2022年底基于人工智能技术的软件版本定出的。

对于“无目标病理”组的患者来说, 对BI-RADS分级的动态没有进行评估, 这可以被认为是

研究的另一个局限性。

结论

本研究显示了使用人工智能技术对乳房X光照片进行首次描述的可行性和发展前途。一种医疗器械, 即基于人工智能技术的软件, 具有与放射科医生对乳房X光照片进行两次查看后获得的灵敏度和特异性一致或超过该数值的灵敏度和特异性。基于人工智能技术的软件的首次描述与放射科医生的第二次描述相结合的应用模式, 使全国范围内的经济效益每年达到6亿至55亿卢布。

ADDITIONAL INFORMATION

Funding source. This article was prepared by a group of authors as a part of the research and development effort titled “Evidence-based methodologies for sustainable development of artificial intelligence in medical imaging”, registration number to EGISU: 123031500004-5.

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

Authors' contribution. All authors made a substantial contribution to the conception of the work, acquisition, analysis, interpretation of data for the work, drafting and revising the work, final approval of the version to be published and agree to be accountable for all aspects of the work. Yu.A. Vasilev, I.A. Tyrov — research concept, final proofreading of the text; A.V. Vladzimirsky — research concept and design, writing and editing of the text; K.M. Arzamasov — research concept and design, analysis of the data obtained, writing and editing of the text; I.M. Shulkin — research design, collection and processing of materials, final proofreading of the text; D.D. Kozhikhina — research design, collection and processing of materials; L.D. Pestrenin — analysis of the data obtained, writing the text.

СПИСОК ЛИТЕРАТУРЫ

1. Злокачественные новообразования в России в 2021 году (заболеваемость и смертность) / под ред. А.Д. Каприна, В.В. Старинского, А.О. Шахзадовой. Москва, 2022. 252 с.
2. Состояние онкологической помощи населению России в 2021 году / под ред. А.Д. Каприна, В.В. Старинского, А.О. Шахзадовой. Москва, 2022. 239 с.
3. Chen Y., James J.J., Michalopoulou E., et al. Performance of radiologists and radiographers in double reading mammograms: The UK national health service breast screening program // *Radiology*. 2023. Vol. 306, N 1. P. 102–109. doi: 10.1148/radiol.212951
4. Euler-Chelpin M.V., Lillholm M., Napolitano G., et al. Screening mammography: Benefit of double reading by breast density // *Breast Cancer Res Treat*. 2018. Vol. 171, N 3. P. 767–776. doi: 10.1007/s10549-018-4864-1
5. Hickman S.E., Woitek R., Le E.P., et al. Machine learning for workflow applications in screening mammography: Systematic review and meta-analysis // *Radiology*. 2022. Vol. 302, N 1. P. 88–104. doi: 10.1148/radiol.2021210391
6. Liu J., Lei J., Ou Y., et al. Mammography diagnosis of breast cancer screening through machine learning: A systematic review and meta-analysis // *Clin Exp Med*. 2022. doi: 10.1007/s10238-022-00895-0
7. Рожкова Н.И., Ройтберг П.Г., Варфоломеева А.А., и др. Сегментационная модель скрининга рака молочной железы на основе нейросетевого анализа рентгеновских изображений // *Сеченовский вестник*. 2020. Т. 11, № 3. С. 4–14. doi: 10.47093/2218-7332.2020.11.3.4-14
8. Компьютерное зрение в лучевой диагностике: первый этап Московского эксперимента. Монография / под ред. Ю.А. Васильева, А.В. Владимировского. Москва: Издательские решения, 2022. 388 с.
9. Патент РФ на изобретение № 2022617324/05.04.2022. Бюл. № 4. Морозов С.П., Андрейченко А.Е., Четвериков С.Ф., и др. Веб-инструмент для выполнения ROC-анализа результатов диагностических тестов. Режим доступа: <https://www.elibrary.ru/item.asp?id=48373757>. Дата обращения: 10.03.2023.

10. Morozov S.P., Vladzimirsky A.V., Klyashtorny V.G., et al. Clinical acceptance of software based on artificial intelligence technologies (radiology). Moscow, 2019. 45 p. (Ser. Best practices in medical imaging).
11. Schaffter T., Buist D.S., Lee C.I., et al. Evaluation of combined artificial intelligence and radiologist assessment to interpret screening mammograms // *JAMA Netw Open*. 2020. Vol. 3, N 3. P. e200265. doi: 10.1001/jamanetworkopen.2020.0265

12. Wan Y., Tong Y., Liu Y., et al. Evaluation of the combination of artificial intelligence and radiologist assessments to interpret malignant architectural distortion on mammography // *Front Oncol*. 2022. N 12. P. 880150. doi: 10.3389/fonc.2022.880150
13. Leibig C., Brehmer M., Bunk S., et al. Combining the strengths of radiologists and AI for breast cancer screening: A retrospective analysis // *Lancet Digit Health*. 2022. Vol. 4, N 7. P. e507–e519. doi: 10.1016/S2589-7500(22)00070-X

REFERENCES

1. Malignant neoplasms in Russia in 2021 (morbidity and mortality). Ed by A.D. Kaprin, V.V. Starinsky, A.O. Shahzadova. Moscow; 2022. 252 p. (In Russ).
2. The state of oncological assistance to the population of Russia in 2021. Ed by A.D. Kaprin, V.V. Starinsky, A.O. Shahzadova. Moscow; 2022. 239 p. (In Russ).
3. Chen Y, James JJ, Michalopoulou E, et al. Performance of radiologists and radiographers in double reading mammograms: The UK national health service breast screening program. *Radiology*. 2023;306(1):102–109. doi: 10.1148/radiol.212951
4. Euler-Chelpin MV, Lillholm M, Napolitano G, et al. Screening mammography: Benefit of double reading by breast density. *Breast Cancer Res Treat*. 2018;171(3):767–776. doi: 10.1007/s10549-018-4864-1
5. Hickman SE, Woitek R, Le EP, et al. Machine learning for workflow applications in screening mammography: Systematic review and meta-analysis. *Radiology*. 2022;302(1):88–104. doi: 10.1148/radiol.2021210391
6. Liu J, Lei J, Ou Y, et al. Mammography diagnosis of breast cancer screening through machine learning: A systematic review and meta-analysis. *Clin Exp Med*. 2022. doi: 10.1007/s10238-022-00895-0
7. Rozhkova NI, Rojtbjerg PG, Varfolomeeva AA, et al. Neural network-based segmentation model for breast cancer X-ray screening. *Sechenov medical journal*. 2020;11(3):4–14 (In Russ). doi: 10.47093/2218-7332.2020.11.3.4-14
8. Vasilev JA, Vladzimirskyy AV. Computer vision in radiology: The first stage of the Moscow experiment: Monograph. Moscow: Izdatel'skie resheniya; 2022. 388 p. (In Russ).
9. Patent RUS № 2022617324/05.04.2022. Byul. № 4. Morozov SP, Andreichenko AE, Chetverikov SF, et al. A web-based tool for performing ROC analysis of diagnostic test results. Available from: <https://www.elibrary.ru/item.asp?id=48373757>. Accessed: 10.03.2023. (In Russ).
10. Morozov SP, Vladzimirsky AV, Klyashtorny VG, et al. Clinical acceptance of software based on artificial intelligence technologies (radiology). Moscow; 2019. 45 p. (Ser. Best practices in medical imaging).
11. Schaffter T, Buist DS, Lee CI, et al. Evaluation of combined artificial intelligence and radiologist assessment to interpret screening mammograms. *JAMA Netw Open*. 2020;3(3):e200265. doi: 10.1001/jamanetworkopen.2020.0265
12. Wan Y, Tong Y, Liu Y, et al. Evaluation of the combination of artificial intelligence and radiologist assessments to interpret malignant architectural distortion on mammography. *Front Oncol*. 2022;(12):880150. doi: 10.3389/fonc.2022.880150
13. Leibig C, Brehmer M, Bunk S, et al. Combining the strengths of radiologists and AI for breast cancer screening: A retrospective analysis. *Lancet Digit Health*. 2022;4(7):e507–e519. doi: 10.1016/S2589-7500(22)00070-X

AUTHORS' INFO

* **Lev D. Pestrenin**, Junior Research Associate;
address: 24/1 Petrovka street, 127051 Moscow, Russia;
ORCID: 0000-0002-1786-4329;
eLibrary SPIN: 7193-7706;
e-mail: PestreninLD@zdrav.mos.ru

Yuriy A. Vasilev, MD, Cand. Sci. (Med);
ORCID: 0000-0002-0208-5218;
eLibrary SPIN: 4458-5608;
e-mail: npcmmr@zdrav.mos.ru

Ilya A. Tyrov;
ORCID: 0000-0001-9337-624X;
eLibrary SPIN: 8625-3458;
e-mail: npcmmr@zdrav.mos.ru

Anton V. Vladzimirskyy, MD, Dr. Sci. (Med), Professor;
ORCID: 0000-0002-2990-7736;
eLibrary SPIN: 3602-7120;
e-mail: npcmmr@zdrav.mos.ru

ОБ АВТОРАХ

* **Пестренин Лев Дмитриевич**, мл. науч. сотр.;
адрес: Россия, 127051, Москва, ул. Петровка, д. 24, стр. 1;
ORCID: 0000-0002-1786-4329;
eLibrary SPIN: 7193-7706;
e-mail: PestreninLD@zdrav.mos.ru

Васильев Юрий Александрович, канд. мед. наук;
ORCID: 0000-0002-0208-5218;
eLibrary SPIN: 4458-5608;
e-mail: npcmmr@zdrav.mos.ru

Тыров Илья Александрович;
ORCID: 0000-0001-9337-624X;
eLibrary SPIN: 8625-3458;
e-mail: npcmmr@zdrav.mos.ru

Владзими́рский Анто́н Вячесла́вович, д-р мед. наук, профессор;
ORCID: 0000-0002-2990-7736;
eLibrary SPIN: 3602-7120;
e-mail: npcmmr@zdrav.mos.ru

* Corresponding author / Автор, ответственный за переписку

Kirill M. Arzamasov, MD, Cand. Sci. (Med);
ORCID: 0000-0001-7786-0349;
eLibrary SPIN: 3160-8062;
e-mail: ArzamasovKM@zdrav.mos.ru

Igor M. Shulkin;
ORCID: 0000-0002-7613-5273;
eLibrary SPIN: 5266-0618;
e-mail: npcmr@zdrav.mos.ru

Daria D. Kozhikhina;
ORCID: 0000-0001-7690-8427;
eLibrary SPIN: 5869-3854;
e-mail: npcmr@zdrav.mos.ru

Арзамасов Кирилл Михайлович, канд. мед. наук;
ORCID: 0000-0001-7786-0349;
eLibrary SPIN: 3160-8062;
e-mail: ArzamasovKM@zdrav.mos.ru

Шулькин Игорь Михайлович;
ORCID: 0000-0002-7613-5273;
eLibrary SPIN: 5266-0618;
e-mail: npcmr@zdrav.mos.r

Кожихина Дарья Дмитриевна;
ORCID: 0000-0001-7690-8427;
eLibrary SPIN: 5869-3854;
e-mail: npcmr@zdrav.mos.ru