

DOI: <https://doi.org/10.17816/DD501759>

Технологические дефекты программного обеспечения с искусственным интеллектом

В.В. Зинченко, К.М. Арзамасов, Е.И. Кремнева, А.В. Владзимирский, Ю.А. Васильев

Научно-практический клинический центр диагностики и телемедицины, Москва, Российская Федерация

АННОТАЦИЯ

Обоснование. Технологические дефекты в работе программного обеспечения с искусственным интеллектом являются критически важными при принятии решения о практической применимости и клинической ценности программного обеспечения с искусственным интеллектом.

Цель — анализ и систематизация технологических дефектов, возникающих при работе программного обеспечения с искусственным интеллектом для анализа медицинских изображений.

Материалы и методы. В рамках эксперимента по использованию инновационных технологий в области компьютерного зрения для анализа медицинских изображений и дальнейшего применения в системе здравоохранения города Москвы проводится мониторинг технологических параметров для всех участвующих решений как на этапе апробации, так и на этапе опытной эксплуатации. В статье представлена графическая информация о среднем числе технологических дефектов для профилактического направления, модальность «Маммография», за 2021 год. Этот период выбран как наиболее показательный, характеризующийся активным развитием программного обеспечения с искусственным интеллектом с позиции увеличения технической стабильности их работы. С целью оценки применимости подхода по выявлению технологических дефектов аналогичный анализ проводился для направления обнаружения внутрисерепных кровоизлияний на компьютерных томограммах головного мозга за 2022–2023 годы.

Результаты. В ходе исследования было проанализировано программное обеспечение с искусственным интеллектом по модальностям «Маммография» (2 алгоритма) и «Компьютерная томография головного мозга» (1). Всего для модальности «Маммография» собрано 14 выборок по 20 исследований; для модальности «Компьютерная томография» — 12 выборок по 80 исследований. Для каждого типа дефекта были построены графики, а для каждой из модальностей были построены линии тренда. Коэффициенты уравнений линий трендов указывают на тенденцию к снижению числа технологических дефектов.

Заключение. Проведённый анализ позволяет проследить тенденцию к снижению числа технологических дефектов, что может свидетельствовать о доработке программного обеспечения с искусственным интеллектом и повышении его качества благодаря периодическому мониторингу. Кроме того, такой результат показывает универсальность использования как для профилактических методов, так и для экстренных.

Ключевые слова: искусственный интеллект; программное обеспечение с искусственным интеллектом; технологический мониторинг; технологические дефекты; Московский эксперимент.

Как цитировать:

Зинченко В.В., Арзамасов К.М., Кремнева Е.И., Владзимирский А.В., Васильев Ю.А. Технологические дефекты программного обеспечения с искусственным интеллектом // Digital Diagnostics. 2023. Т. 4, № 4 С. 593–604. DOI: <https://doi.org/10.17816/DD501759>

DOI: <https://doi.org/10.17816/DD501759>

Technological defects in software based on artificial intelligence

Viktoriya V. Zinchenko, Kirill M. Arzamasov, Elena I. Kremneva, Anton V. Vladzemyrskyy, Yuriy A. Vasilev

Scientific and Practical Clinical Center for Diagnostics and Telemedicine Technologies, Moscow, Russian Federation

ABSTRACT

BACKGROUND: Technological defects in the use of artificial intelligence software are critical when deciding on the practical applicability and clinical value of artificial intelligence software.

AIM: To conduct an analysis and systematization of technological defects occurring when artificial intelligence software analyzes medical images.

MATERIALS AND METHODS: As part of the experiment on the use of innovative computer vision technologies for the analysis of medical images and further application in the Moscow healthcare system, technological parameters of all artificial intelligence software are monitored at the testing and operation stages of the trial. This article presents graphical information on the average number of technological defects in mass mammography screening in 2021. This period was chosen as the most indicative and characterized by the active development of artificial intelligence software and increased technical stability of its performance. To assess the applicability of the analysis for technological defects, a similar analysis was conducted for the direction of detection of intracranial hemorrhage on computed tomography scans of the brain for 2022–2023.

RESULTS: During the study, artificial intelligence software used for mammography (two algorithms) and brain computed tomography (one algorithm) were analyzed. Fourteen mammography samples were collected for technological monitoring during the identified period, each from 20 studies, and 12 brain computed tomography samples were obtained, each from 80 studies. Graphs were constructed for each type of defect, and trend lines were plotted for each modality. The coefficients of the trend line equations indicate a downward tendency in the number of technological defects.

CONCLUSION: This analysis allows tracing a downward trend in the number of technological defects, which may indicate a refinement of artificial intelligence software and an increase in its quality because of periodic monitoring. It also shows the versatility of use for both preventive and emergency methods.

Keywords: artificial intelligence; artificial intelligence software; technological monitoring; technological defects; Moscow experiment.

To cite this article:

Zinchenko VV, Arzamasov KM, Kremneva EI, Vladzemyrskyy AV, Vasilev YuA. Technological defects in software based on artificial intelligence. *Digital Diagnostics*. 2023;4(4):593–604. DOI: <https://doi.org/10.17816/DD501759>

Received: 20.06.2023

Accepted: 21.11.2023

Published online: 06.12.2023

DOI: <https://doi.org/10.17816/DD501759>

人工智能软件的技术缺陷

Viktoria V. Zinchenko, Kirill M. Arzamasov, Elena I. Kremneva, Anton V. Vladzmyrskyy, Yuriy A. Vasilev

Scientific and Practical Clinical Center for Diagnostics and Telemedicine Technologies, Moscow, Russian Federation

简评

论证。人工智能软件性能方面的技术缺陷是确定人工智能软件实用性和临床价值的关键。

该研究的目的是对医学影像分析人工智能软件运行中的技术缺陷进行分析并使之系统化。

材料和方法。在莫斯科市进行了一项《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》。在实验框架内，对所有参与解决方案的技术参数进行监测。监测是在批准阶段和试运行阶段进行的。本文以图表形式介绍2021年“乳房摄影术”预防方向的平均技术缺陷数量。这一时期被选为最有意义的时期。这一时期的特点是从提高操作技术稳定性的角度出发，积极开发人工智能软件。为了评估该方法在发现技术缺陷方面的适用性，我们对2022–2023年脑部CT扫描颅内出血的检测方向进行了类似的分析。

结果。本研究分析了“乳房摄影术”（2种算法）和“脑计算机断层扫描”（1种）模式的人工智能软件。在“乳房X射线照相术”模式中，共收集了14个样本，共有20项研究。在“脑计算机断层扫描”模式中，共收集了12个样本，共有80项研究。我们对每种缺陷类型都绘制了图表，对每种模式绘制了趋势线。趋势线公式的系数表明了，技术缺陷的数量呈下降趋势。

结论。通过分析，我们发现了减少技术缺陷数量的趋势。这可能表明人工智能软件的完善，以及通过定期监测，软件质量的提升。此外，这一结果还显示使用预防和应急方法的通用性。

关键词：人工智能；人工智能软件；技术监测；技术缺陷；莫斯科实验。

引用本文：

Zinchenko VV, Arzamasov KM, Kremneva EI, Vladzmyrskyy AV, Vasilev YuA. 人工智能软件的技术缺陷. *Digital Diagnostics*. 2023;4(4):593–604. DOI: <https://doi.org/10.17816/DD501759>

收到: 20.06.2023

接受: 21.11.2023

发布日期: 06.12.2023

论证

人工智能技术软件可以帮助医务人员完成常规和复杂的任务,并提高为患者提供医疗服务的水平、可用性和速度[1-3]。这在很大程度上要归功于《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“实验”),以及国外和国内在医疗领域使用人工智能工作经验的继承性[4-7]。该实验的目的是根据先进新技术的数据分析结果,科学研究在莫斯科医疗系统中使用医疗决策支持方法的可能性。为21个放射诊断领域的人工智能工作结果制定了要求。目前,已有50多项基于人工智能的决策结果可供医生使用。截至2023年9月底,已处理超过1000万个检查。

新技术在医疗保健中的应用需要强制遵守安全规则,因此人工智能软件的开发、调试和应用阶段必须受到强制性控制[8]。人工智能软件在运行过程中需要特别控制,这也是因为它被用于培训人群以外的人群时,结果可能会出现偏差[9, 10]。

实验中使用了一系列测试来控制人工智能软件检查处理的质量。自测试是初始阶段,旨在了解人工智能软件与因处理而提交的检查(输入数据)之间的技术兼容性。下一阶段是功能测试,

确定人工智能软件所声明功能的可用性,包括其可操作性。人工智能软件要从技术和临床两个角度进行评估,即由技术专家和医学专家进行评估。校准测试是确定人工智能软件性能指标的阶段,主要指标是ROC曲线下面积。

在成功完成所有测试的情况下,允许使用人工智能软件,同时根据工作结果对算法进行技术和临床监测。按照国际研究,技术测试(技术参数监测)是产品验证不可分割的一部分,产品验证全面测试时进行,以便在实际临床实践中使用[11]。因此,我们在这项研究的框架内重点关注技术缺陷的监测。

目的

从实验结果中研究人工智能软件的技术缺陷,对其进行分析、统计,以及确定其对临床实践中人工智能软件的安全和质量的影响。

材料和方法

执行条件

根据表1(左列,按照莫斯科市卫生局于2021年1月26日第51号命令)所示的缺陷类别,对2021

表1. 莫斯科市卫生局命令中技术缺陷标准的相关性

莫斯科市卫生局于2021年1月26日第51号命令规定的技术缺陷 (与文章中提供的乳腺X线摄影数据相关)	莫斯科市卫生局于2022年11月3日第160号命令规定的技术缺陷(重组) (与文章中提供的大脑CT数据相关)
a组: 分析一个检查的时间超过6.5分钟	a组: 分析一个检查的时间超过6.5分钟
b组: 无分析检查结果	b组: 无分析检查结果
-	c组: 使用人工智能的软件功能操作不正确,妨碍了放射科医生的工作或使其无法高质量地工作
d2: 无附加系列	c1: 无附加系列
d3: 无DICOM SR	c2: 无DICOM SR
d4: 存在两个或多个DICOM SR	c3: 存在两个或多个DICOM SR
d5: 无人工智能软件名称	c4: 无人工智能软件名称
d6: 无人工智能软件版本信息	c5: 无人工智能软件版本信息
-	d组: 与显示图像区域有关的缺陷
c1: 图像被裁剪	d1: 附加系列中的图像被裁剪
c2: 亮度/对比度发生变化	d2: 附加系列中的亮度/对比度与原始图像不符
c3: 未对所有相关图像进行分析	d3: 未对所有相关图像进行分析
d1: 完全没有使用人工智能软件的工作结果	被排除在外
d7: 无“仅供研究/科学之用”的警告标志	d4: 无“仅供研究/科学之用”的警告标志
d8: 无病理标记	f: 与临床工作有关的缺陷
E1: DICOM SR 和补充系列信息之间的矛盾	被排除在外
F: 修改原始检查系列	d5: 修改原始检查系列
-	e类: 其他破坏检查结果文件完整性和内容的缺陷,包括导致诊断解释受限的缺陷在内
E2: 在目标器官外进行标记	e1: 在目标器官外进行标记
E3: 分析的解剖区域、投影或系列不正确	e2: 分析的解剖区域、投影或系列不正确

注: SR - structure report.

年报告期内“乳腺X线摄影”模式的人工智能软件所分析的检查进行了监测[12]:

- a组: 分析一个检查的时间超过6.5分钟。该时间限制是根据经验得出的,即人工智能软件描述检查结果以便放射科医生使用其结果所需的平均时间。
- b组: 无分析检查结果。
- c组: 人工智能软件结果中包含的图像与原生(原始)检查图像不一致(失真)。在极少数情况下,元数据修改可能会导致查看检查时的设置变化,从而严重影响原始图像的可视化。
- d组: 人工智能软件声明的功能操作不当,导致医生难以或无法高质量地完成工作,包括裁剪图像、改变亮度/对比度、缺少结果描述、缺少病理标记。
- e组: 其他破坏检查结果文件完整性和内容的缺陷,导致诊断解释受限,包括在目标器官外进行标记、人工智能软件分析不正确的解剖区域。
- f组: 修改原始检查系列。2022年,对缺陷进行了重组,在处理“CT”模式的监测数据时考虑到了这一点(见表1右列)。

研究期限

作为实验的一部分,监测每月进行一次,直到人工智能软件的参与结束。监测报告期为一个日历月。每月10日和20日形成a组和b组缺陷监测的中期报告,并发送给人工智能软件的制造商。

对于乳腺X线摄影样本,文章提供了2021年3月至12月的缺陷信息;对于大脑CT样本,文章提供了2022年5月至2023年5月的缺陷信息。对于不同的人工智能软件,由于进入实验的时刻和收到反馈后的改进时间不同,监测周期性也不同。

技术监测由一个专家组进行,该专家组包括技术人员和放射科医生,他们接受了额外的进行监测培训和使用特定人工智能软件的指导。此外,还制定了统一的内部报告表和技术监测说明,用于报告监测情况。

统计分析

在进行技术监测的过程中,使用了伪随机抽取的数据集(检查样本),其中考虑到了以下比例:25%的检查中人工智能软件未检测到病理(“无病理”组),75%的检查中检测到病理(“有病理”组)。对人工智能软件得出结果的检查进行了技术缺陷检查。如果超过测试期间设定的最佳阈值,该检查被定义为“有病理”组;否则被归类为“无病理”组[13, 14]。

在2021年的实验中,伪随机样本规模为每月20个检查。需要注意的是,这仍然是该项目的试点阶段,根据诺模图,在统计显著性水平为0.05时,功率水平为42.5%。样本元素之间的标准差为0.79[15]。在2021年后的正式项目中,采用风险分析方法,样本量为80个检查(S. F. Chetverikov

等人的文章[13]中得到证实)。这与大脑CT检查的每月样本量(80个检查)完全相同。

结果

为了对“乳腺X线摄影”模式进行技术监测,共使用了包括20个检查在内的14个样本。在2021年3月至12月期间,每月将生成的伪随机样本发送给所有正在使用(非开发中)的人工智能软件进行测试。

为了评估该方法在检测技术缺陷方面的适用性,我们对检测颅内出血的“大脑CT”模式生成的伪随机样本进行了类似的分析。从2022年5月至2023年5月,每月向人工智能软件发送80个检查(共12个样本,每个样本80个检查)进行测试。

为了绘制2021年3月至12月(“乳腺X线摄影”模式)或2022年5月至2023年5月(“CT”模式)期间的技术缺陷动态变化图,使用了人工智能软件的所有技术监测结果的总统计数据。技术缺陷的数量视为数据集中检查总数的百分比。

图1给出2021年3月至2021年12月期间“乳腺X线摄影”模式技术缺陷平均数量的动态变化,其中纵轴为存在缺陷的情况(以占样本检查总数的百分比表示),横轴为以月为单位的报告期。图2给出“大脑CT模式”(2022年5月至2023年5月)的类似信息。

对于“乳腺X线摄影”模式,缺陷情况按表1左列列出。图1显示,在研究监测期开始时,c组(与图像区域显示有关的缺陷)、d组(人工智能软件所声明功能的不正确操作)和b组(无分析检查结果)缺陷占多数。到研究期结束时,只剩下c组缺陷,但其所占比例也明显下降。

对于“CT”模式,缺陷情况按表1右列列出。图2显示,对于“大脑CT”模式颅内出血检测中,除b组(无分析检查结果)外,缺陷与所有缺陷的样本的比例低于其他分析模式。d组(与图像区域显示有关的缺陷)和e组(文件完整性和内容的缺陷)的缺陷数量有所下降(百分比),而b组的缺陷在月与月之间变化很大。

为了量化这一趋势,我们添加了相应的趋势线。这些趋势线是 $k \times x + b$ 形式的线性函数,其中系数 k 定义近似曲线的斜率,即表示缺陷数量增加或减少的趋势,而系数 b 则对应于监测开始时的缺陷数量。对所有人工智能软件进行了特定模式的近似,同时对整个数据集进行了近似(见图1-2)。知道了系数 k ,就可以预测每个人工智能软件的技术缺陷消除动态变化,或整个方向的技术缺陷消除动态变化。

图3-6给出人工智能软件的技术缺陷示例。

讨论

根据获得和分析的研究结果,“乳腺X线摄影”模式在减少技术缺陷数量方面表现出良好的趋势(见图1,趋势线)。尽管存在b组缺陷值,“大脑CT”模式的人工智能软件在减少技术缺陷数量方面的趋

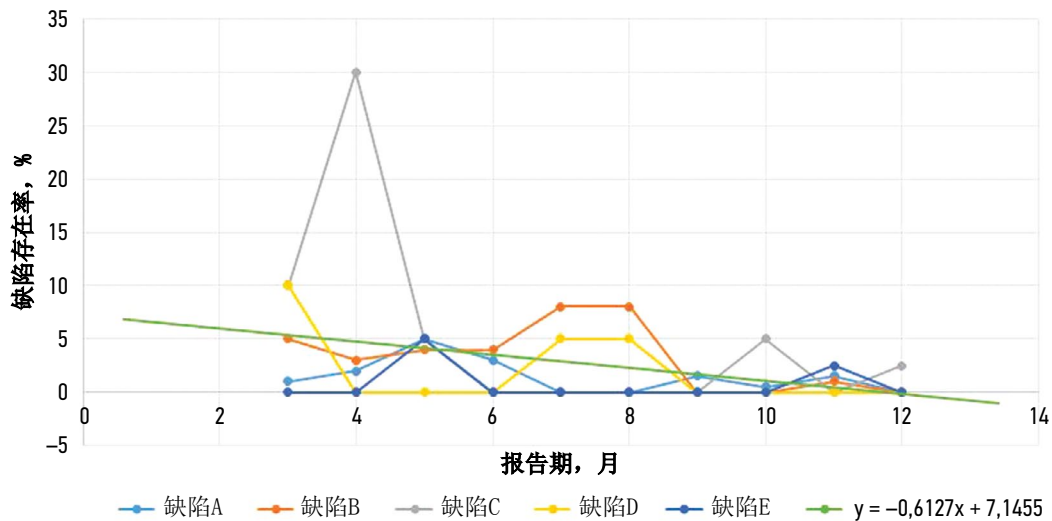


图1。“乳腺X线摄影”模式的人工智能软件检测到的每种技术缺陷的平均数量检测动态变化。根据莫斯科市卫生局于2021年1月26日第51号命令，缺陷按组分布。

势较为平缓(见图2,趋势线)。这反过来又可以解释为,某些技术缺陷的波动与自动检测以及人工智能制造商的快速反馈和及时软件修改(版本更改或错误纠正)有关。

值得注意的是,在2021年9月至10月期间,“乳腺X线摄影”模式的人工智能软件版本发生了变化,与此同时,b组和d组的平均缺陷数量也有所减少(见图1)。这可能表明人工智能软件的成功开发,这反过来又可能表明所介绍的技术监测方法的有效性。

应该指出的是,在技术监测框架内识别技术缺陷是综合测试的重要组成部分,目的是使人工智能软件不仅在辐射诊断领域,而且在整个医学领域更安全、更高质量和更高效地运行。分析结果表明,人工智能软件的质量随着缺陷数量的减少而提高。因此,人工智能软件能获得更多用户的信任,以最少的缺陷数量工作,为医生提供帮助[16, 17]。

技术缺陷重组

根据本文介绍的技术缺陷监测和分析结果,2022年对技术缺陷组进行了重组。“大脑CT”模式的人工智能软件缺陷监测(是否存在颅内出血)是根据更新后的组别分类(见表1右列)进行的。a组和b组的缺陷以自动模式进行检查,而c组、d组和e组的缺陷则需要专家手动检查。

莫斯科市卫生局于2022年11月3日第160号命令中列出了最新的技术缺陷清单,该清单仍具有现实意义[18]。S. F. Chetverikov等人的文章[13]介绍了增加样本检查数量的理由。对技术缺陷进行这样的重组,使得优化分析人工智能软件监测结果的专家的工作成为可能。

此外,根据在实验条件下对人工智能软件的技术监测结果,按照2021年关于“乳腺X线摄影”模式的命令,在人工智能软件作为医疗设备的安全性方面,技术缺陷可有条件地分为三组:

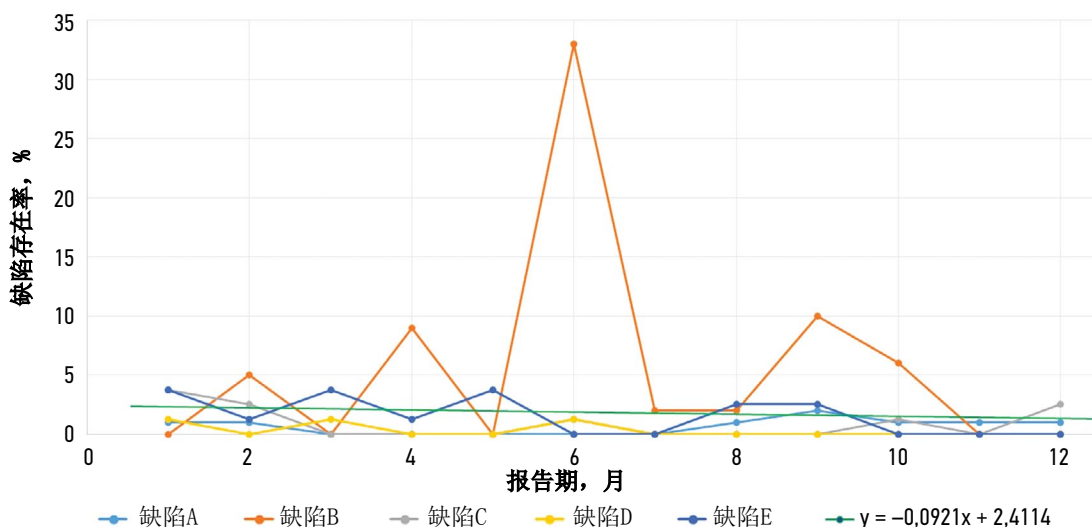


图2。“脑部CT”模式的人工智能软件检测到的每种技术缺陷的平均数量检测动态变化(是否存在颅内出血)。根据莫斯科市卫生局于2022年11月3日第160号命令,缺陷按组分布。

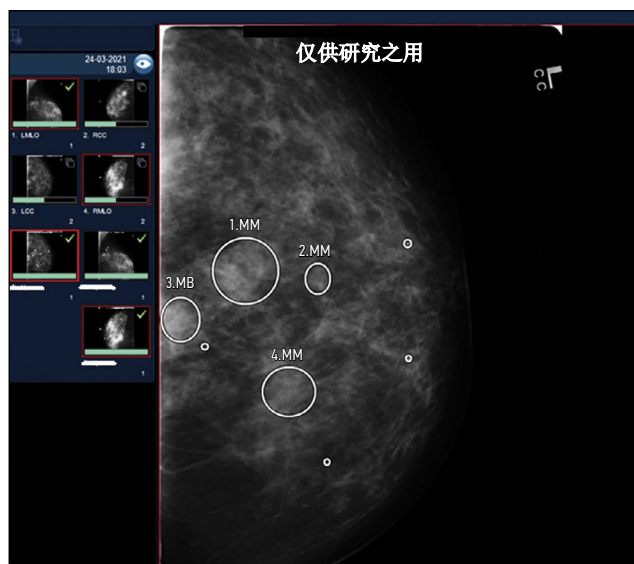


图3. 缺陷：未对所有相关图像进行分析。“乳腺X线摄影”模式。

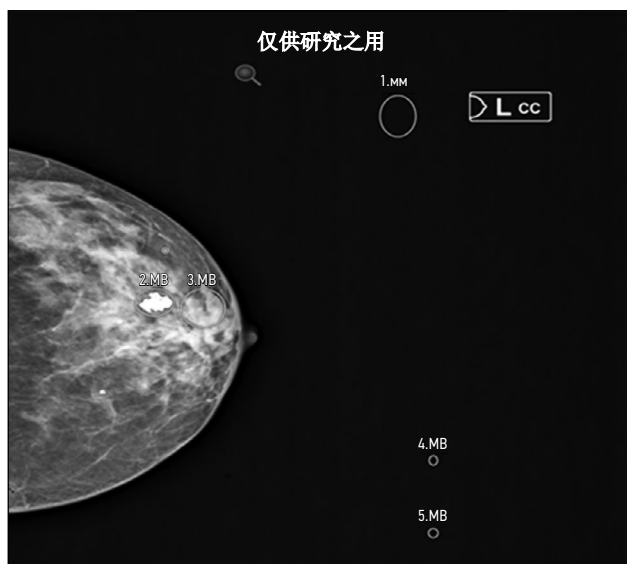


图4. 缺陷：在目标器官外进行标记。“乳腺X线摄影”模式。

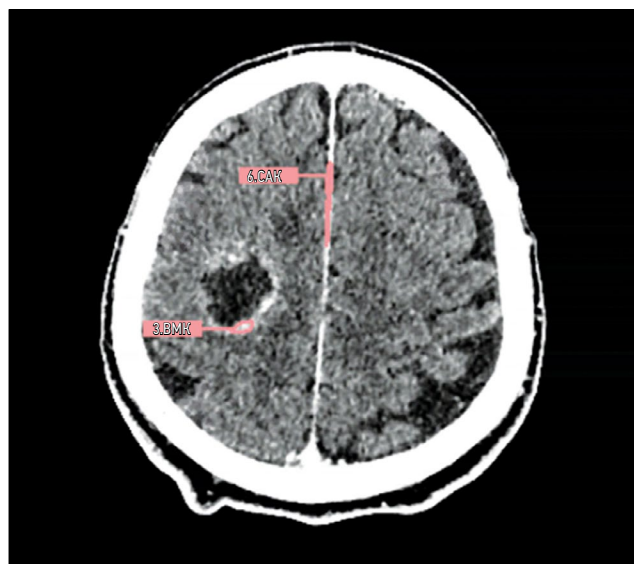


图5. 缺陷：分析的系列不正确（对比后CT而非原始CT）。“CT”模式。

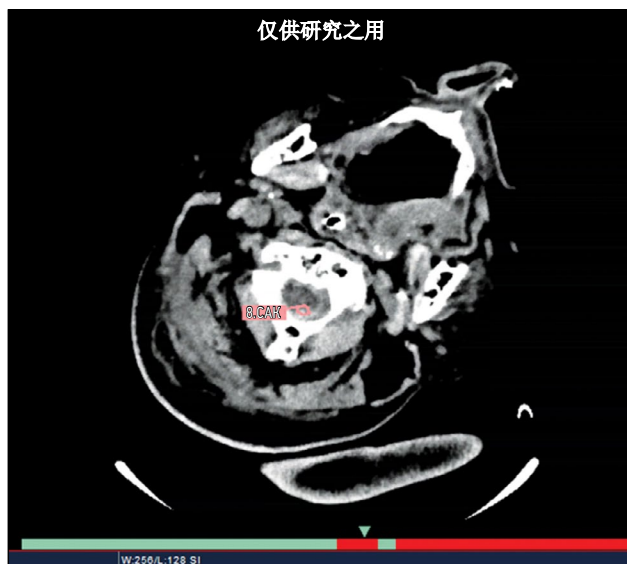


图6. 缺陷：在目标器官外进行标记（对比后CT而非原始CT）。“CT”模式。

- **影响患者安全和医生工作：** 未能实现制造商声明的功能；影响放射科医生或妨碍其工作的备注；原始检查数据不可逆转的失真。例如，这组缺陷包括d组缺陷（d2、d3、d4），以及f组缺陷（改变检查的原始系列，直接影响使用人工智能软件的安全性）。需要单独考虑d7缺陷（没有警告标志：“仅供研究/科学之用”）。该缺陷仅可能在科学研究框架内出现，但绝不可能在将人工智能软件作为医疗设备使用的框架内出现。
- **不影响患者安全，但影响医生工作：** 不符合普遍接受的呈现检查解释结果的标准的功能缺陷。这类缺陷包括e组和c组（c1、c2、c3）。
- **不影响患者安全，也不影响医生工作：** 为使医生工作更方便、更直观、操作性更强而需

要消除的小缺陷。这类缺陷包括d5、d6和d8缺陷。

对于“CT”模式，由于缺陷重组（表1），从2023年11月起，根据2021年的命令，在安全性方面分为三组：

- **影响患者安全和医生工作：** c组缺陷（c1、c2、c3）以及d4、d5缺陷。
- **不影响患者安全，但影响医生工作：** e组和d组（d1、d2、d3）缺陷。
- **不影响患者安全，也不影响医生工作：** c4和c5缺陷。

图7 - 8给出两种模式按组别和月份划分的缺陷数量变化情况。

就“乳腺X线摄影”模式而言（图7），由于人工智能软件改进，影响患者安全和医生工作的缺

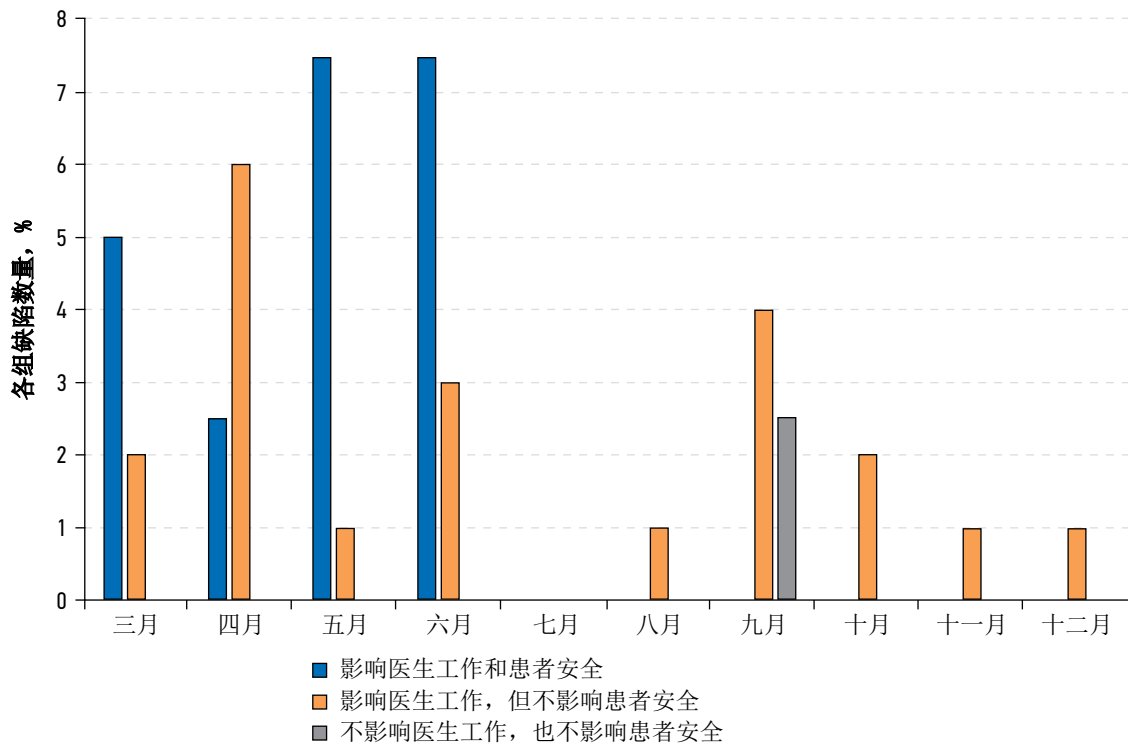


图7. “乳腺X线摄影”模式按组别和月份划分的缺陷数量变化情况。

陷在6月之后就不再被发现。到研究期结束时，影响医生工作，但不影响病人安全的缺陷也有下降趋势。

就“大脑CT”模式而言，最常出现的缺陷（影响医生工作，但不影响患者安全）并没有明显的下降趋势。

本文介绍的技术缺陷监测方法允许实现对算法技术稳定性的控制，这对于评估人工智能软件及其安全性具有重要的实际意义。正如我们在大脑CT分析数据的例子中看到的那样，在流程上控制人工智能软件运行的应用方法使我们能够发现技术缺陷并完善解决方案，最终提高了人工智能

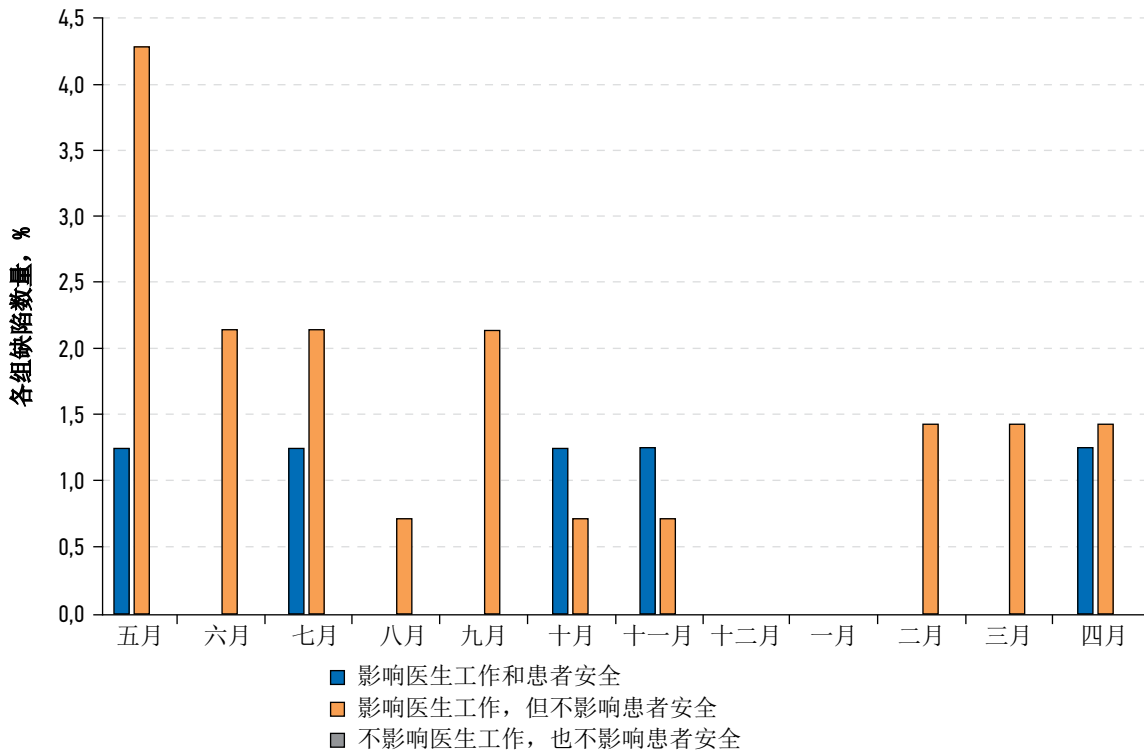


图8. “CT”模式按组别和月份划分的缺陷数量变化情况。

软件的技术稳定性。因此，所开发的方法被证明是提高人工智能软件技术稳定性的有效和通用工具。

结论

本文列出了人工智能软件实施过程中出现的主要技术缺陷，并介绍了基于定期随机对照测试的技术缺陷监控方法，以提高人工智能软件的技术稳定性。在识别技术缺陷时开发的测试人工智能软件的方法似乎是在实际临床实践中工作条件下测试人工智能软件的安全、质量和有效性监测的一部分。

ADDITIONAL INFORMATION

Funding source. Analysis of technological defects in computed tomography dataset with or without intracranial bleeding was funded

СПИСОК ЛИТЕРАТУРЫ

1. Владимирский А.В., Васильев Ю.А., Арзамасов К.М. и др. Компьютерное зрение в лучевой диагностике: первый этап Московского эксперимента / под ред. Ю.А. Васильева, А.В. Владимирского. Издательские решения, 2022.
2. Ranschaert E.R., Morozov S., Algra P.R., editors. *Artificial Intelligence in Medical Imaging*. Berlin : Springer, 2019. doi: 10.1007/978-3-319-94878-2
3. Гусев А.В., Добридюк С.Л. Искусственный интеллект в медицине и здравоохранении // Информационное общество. 2017. № 4-5. С. 78–93.
4. Шутов Д.В., Шарова Д.Е., Абуладзе Л.П., Дроздов Д.В. Системы искусственного интеллекта в клинической физиологии: как сделать их обучение эффективным? // *Digital Diagnostics*. 2023. Т. 4, № 1. С. 81–88. doi: 10.17816/DD123559
5. Мелдо А.А., Уткин Л.В., Трофимова Т.Н. Искусственный интеллект в медицине: современное состояние и основные направления развития интеллектуальной диагностики // *Лучевая диагностика и терапия*. 2020. Т. 11, № 1. С. 9–17. doi: 10.22328/2079-5343-2020-11-1-9-17
6. Recht M.P., Dewey M., Dreyer K., et al. Integrating artificial intelligence into the clinical practice of radiology: challenges and recommendations // *European radiology*. 2020. Vol. 30, N 6. P. 3576–3584. doi: 10.1007/s00330-020-06672-5
7. Larson D.B., Harvey H., Rubin D.L., et al. Regulatory Frameworks for Development and Evaluation of Artificial Intelligence-Based Diagnostic Imaging Algorithms: Summary and Recommendations // *Journal of the American College of Radiology*. 2021. Vol. 18, N 3 Pt A. P. 413–424. doi: 10.1016/j.jacr.2020.09.060
8. Zinchenko V., Chetverikov S., Ahmad E., et al. Changes in software as a medical device based on artificial intelligence technologies // *International Journal of Computer Assisted Radiology and Surgery*. 2022. Vol. 17. P. 1969–1977. doi: 10.1007/s11548-022-02669-1
9. Nomura Y., Miki S., Hayashi N., et al. Novel platform for development, training, and validation of computer-assisted detection/diagnosis software // *International Journal of Computer Assisted Radiology and Surgery*. 2020. Vol. 15, N 4. P. 661–672. doi: 10.1007/s11548-020-02132-z

by Russian Science Foundation Grant № 22-25-20231, <https://rscf.ru/project/22-25-20231/>.

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. The contribution is distributed as follows: V.V. Zinchenko — structuring and analysis of the results obtained (mammography modality), writing the manuscript of the article; K.M. Arzamasov — obtaining technological monitoring data, analyzing the results obtained, correcting the manuscript of the article; E.I. Kremneva — structuring and analysis of the results obtained (computed tomography modality), writing the manuscript of the article; A.V. Vladzimirskyy — review of the manuscript of the article, formation of the research hypothesis; Yu.A. Vasilev — formation of the research hypothesis, general management of the research.

10. Методические рекомендации по порядку проведения экспертизы качества, эффективности и безопасности медицинских изделий (в части программного обеспечения) для государственной регистрации в рамках национальной системы ФГБУ «ВНИИ-ИМТ» Росздравнадзора. Москва, 2021.

11. Pemberton H.G., Zaki L.A.M., Goodkin O., et al. Technical and clinical validation of commercial automated volumetric MRI tools for dementia diagnosis — a systematic review // *Neuroradiology*. 2021. Vol. 63. P. 1773–1789. doi: 10.1007/s00234-021-02746-3

12. Приказ Департамента здравоохранения города Москвы № 51 от 26.01.2021 «Об утверждении порядка и условий проведения эксперимента по использованию инновационных технологий в области компьютерного зрения для анализа медицинских изображений и дальнейшего применения в системе здравоохранения города Москвы в 2021 году»

13. Четвериков С.Ф., Арзамасов К.М., Андрейченко А.Е., и др. Подходы к формированию выборки для контроля качества работы систем искусственного интеллекта в медико-биологических исследованиях // *Современные технологии в медицине*. 2023. Т. 15, № 2. С. 19. doi: 10.17691/stm2023.15.2.02

14. Зинченко В.В., Арзамасов К.М., Четвериков С.Ф., и др. Методология проведения пострегистрационного клинического мониторинга для программного обеспечения с применением технологий искусственного интеллекта // *Современные технологии в медицине*. 2022. Т. 14, № 5. С. 15–25. doi: 10.17691/stm2022.14.5.02

15. Altman D.G. Statistics and ethics in medical research: III How large a sample? // *British medical journal*. 1980. Vol. 281, N 6251. P. 1336. doi: 10.1136/bmj.281.6251.1336

16. Тыров И.А., Васильев Ю.А., Арзамасов К.М., и др. Оценка зрелости технологий искусственного интеллекта для здравоохранения: методология и ее применение на материалах Московского Эксперимента по компьютерному зрению в лучевой диагностике // *Врач и информационные технологии*. 2022. № 4. С. 76–92. doi: 10.25881/18110193_2022_4_76

17. Владимирский А.В., Гусев А.В., Шарова Д.Е., и др. Методика оценки уровня зрелости информационной системы для здравоохранения // *Врач и информационные технологии*. 2022. № 3. С. 68–84. doi: 10.25881/18110193_2022_3_68

18. Приказ Департамента здравоохранения города Москвы от 03.11.2022 № 160 «Об утверждении порядка и условий проведения эксперимента по использованию инновационных технологий

в области компьютерного зрения для анализа медицинских изображений и дальнейшего применения в системе здравоохранения города Москвы в 2022 году».

REFERENCES

- Vladimirskii AV, Vasil'ev YuA, Arzamasov KM, et al. *Computer Vision in Radiologic Diagnostics: the First Stage of Moscow experiment*. Vasil'ev YuA, Vladimirskii AV, editors. Publishing solutions; 2022. (In Russ).
- Ranschaert ER, Morozov S, Algra PR, editors. *Artificial Intelligence in Medical Imaging*. Berlin: Springer; 2019. doi: 10.1007/978-3-319-94878-2
- Gusev AV, Dobridnyuk SL. Artificial intelligence in medicine and healthcare. *Information Society Journal*. 2017;(4-5):78–93. (In Russ).
- Shutov DV, Sharova DE, Abuladze LR, Drozdov DV. Artificial intelligence in clinical physiology: How to improve learning agility. *Digital Diagnostics*. 2023;4(1):81–88. doi: 10.17816/DD123559
- Meldo AA, Utkin LV, Trofimova TN. Artificial intelligence in medicine: current state and main directions of development of the intellectual diagnostics. *Diagnostic radiology and radiotherapy*. 2020;11(1):9–17. doi: 10.22328/2079-5343-2020-11-1-9-17
- Recht MP, Dewey M, Dreyer K, et al. Integrating artificial intelligence into the clinical practice of radiology: challenges and recommendations. *European radiology*. 2020;30(6):3576–3584. doi: 10.1007/s00330-020-06672-5
- Larson DB, Harvey H, Rubin DL, et al. Regulatory Frameworks for Development and Evaluation of Artificial Intelligence-Based Diagnostic Imaging Algorithms: Summary and Recommendations. *Journal of the American College of Radiology*. 2021;18(3 Pt A):413–424. doi: 10.1016/j.jacr.2020.09.060
- Zinchenko V, Chetverikov S, Ahmad E, et al. Changes in software as a medical device based on artificial intelligence technologies. *International Journal of Computer Assisted Radiology and Surgery*. 2022;17:1969–1977. doi: 10.1007/s11548-022-02669-1
- Nomura Y, Miki S, Hayashi N, et al. Novel platform for development, training, and validation of computer-assisted detection/diagnosis software. *International Journal of Computer Assisted Radiology and Surgery*. 2020;15(4):661–672. doi: 10.1007/s11548-020-02132-z
- Methodological recommendations on the procedure for expert examination of quality, efficiency and safety of medical devices (in terms of software) for state registration under the national system FGBU «VNIIMT» Roszdravnadzor. Moscow; 2021. (In Russ).
- Pemberton HG, Zaki LAM, Goodkin O, et al. Technical and clinical validation of commercial automated volumetric MRI tools for dementia diagnosis — a systematic review. *Neuroradiology*. 2021;63:1773–1789. doi: 10.1007/s00234-021-02746-3
- Order of the Moscow City Health Department No. 51 dated 26.01.2021 «On approval of the procedure and conditions for conducting an experiment on the use of innovative technologies in the field of computer vision to analyze medical images and further application in the health care system of the city of Moscow in 2021». (In Russ).
- Chetverikov SF, Arzamasov KM, Andreichenko AE, et al. Approaches to Sampling for Quality Control of Artificial Intelligence in Biomedical Research. *Modern Technologies in Medicine*. 2023;15(2):19. doi: 10.17691/stm2023.15.2.02
- Zinchenko VV, Arzamasov KM, Chetverikov SF, et al. Methodology for Conducting Post-Marketing Surveillance of Software as a Medical Device Based on Artificial Intelligence Technologies. *Modern Technologies in Medicine*. 2022;14(5):15–25. doi: 10.17691/stm2022.14.5.02
- Altman DG. Statistics and ethics in medical research: III How large a sample? *British medical journal*. 1980;281(6251):1336. doi: 10.1136/bmj.281.6251.1336
- Tyrov IA, Vasilyev YuA, Arzamasov KM, et al. Assessment of the maturity of artificial intelligence technologies for healthcare: methodology and its application based on the use of innovative computer vision technologies for medical image analysis and subsequent applicability in the healthcare system of Moscow. *Medical Doctor and IT*. 2022;4(4):76–92. doi: 10.25881/18110193_2022_4_76
- Vladimirsky AV, Gusev AV, Sharova DE, et al. Health Information System Maturity Assessment Methodology. *Medical Doctor and IT*. 2022;3(3):68–84. doi: 10.25881/18110193_2022_3_68
- Order of the Moscow City Health Department No. 160 dated 03.11.2022 «On Approval of the Procedure and Conditions for Conducting an Experiment on the Use of Innovative Technologies in Computer Vision for Analyzing Medical Images and Further Application in the Moscow City Health Care System in 2022». (In Russ).

AUTHORS' INFO

* Viktoria V. Zinchenko;

address: 24-1 Petrovka str., 1127051, Moscow, Russian Federation;
ORCID: 0000-0002-2307-725X;
eLibrary SPIN: 4188-0635;
e-mail: ZinchenkoVV1@zdrav.mos.ru

Kirill M. Arzamasov, MD, Cand. Sci. (Med.);

ORCID: 0000-0001-7786-0349;
eLibrary SPIN: 3160-8062;
e-mail: ArzamasovKM@zdrav.mos.ru

ОБ АВТОРАХ

* Зинченко Виктория Валерьевна;

адрес: Российская Федерация, 127051, Москва, ул. Петровка, д. 24, стр. 1;
ORCID: 0000-0002-2307-725X;
eLibrary SPIN: 4188-0635;
e-mail: ZinchenkoVV1@zdrav.mos.ru

Арзамасов Кирилл Михайлович, канд. мед. наук;

ORCID: 0000-0001-7786-0349;
eLibrary SPIN: 3160-8062;
e-mail: ArzamasovKM@zdrav.mos.ru

* Corresponding author / Автор, ответственный за переписку

Elena I. Kremneva, MD, Cand. Sci. (Med.);
ORCID: 0000-0001-9396-6063;
eLibrary SPIN: 8799-8092;
e-mail: KremnevaEI@zdrav.mos.ru

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

Yuriy A. Vasilev, MD, Cand. Sci. (Med.);
ORCID: 0000-0002-0208-5218;
eLibrary SPIN: 4458-5608;
e-mail: VasilevYA1@zdrav.mos.ru

Кремнева Елена Игоревна, канд. мед. наук;
ORCID: 0000-0001-9396-6063;
eLibrary SPIN: 8799-8092;
e-mail: KremnevaEI@zdrav.mos.ru

Владзimirский Антон Вячеславович, д-р мед. наук;
ORCID: 0000-0002-2990-7736;
eLibrary SPIN: 3602-7120;
e-mail: VladzimirskijAV@zdrav.mos.ru

Васильев Юрий Александрович, канд. мед. наук;
ORCID: 0000-0002-0208-5218;
eLibrary SPIN: 4458-5608;
e-mail: VasilevYA1@zdrav.mos.ru