

DOI: <https://doi.org/10.17816/DD101099>

Возможности и ограничения использования инструментов машинной обработки текстов в лучевой диагностике

Д.Ю. Кокина, В.А. Гомболевский, К.М. Арзамасов, А.Е. Андрейченко, С.П. Морозов

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

АННОТАЦИЯ

Обоснование. В радиологии важную информацию содержат не только медицинские изображения, но и сопровождающие их текстовые описания, создаваемые врачами-рентгенологами. Идентификация протоколов исследований, содержащих определённые данные, и извлечение этих данных может быть полезным в первую очередь для клинических задач, однако, учитывая большой объём таких данных, необходима разработка машинных алгоритмов анализа.

Цель — оценить возможности и ограничения использования инструментов машинной обработки текстов для поиска патологий в протоколах лучевых исследований.

Материалы и методы. Для создания первого прототипа алгоритма автоматического анализа протоколов были выбраны исследования молочных желёз (маммография) и органов грудной клетки (рентгенография, флюорография, компьютерная томография и низкодозная компьютерная томография), выполненные в лечебно-профилактических учреждениях Москвы, которые участвовали в эксперименте по использованию инновационных технологий в области компьютерного зрения для анализа медицинских изображений. Для каждого вида исследований был первоначально составлен словарь ключевых слов, соответствующий наличию или отсутствию целевых патологий. После первичной автоматической разметки протоколов разработанным инструментом производились выборочная оценка и валидация результатов врачом-рентгенологом. Количество протоколов, проанализированных врачом для обучения и валидации алгоритмов, составило 977 для маммографии, 3196 для рентгенографии, 1608 для флюорографии, 4074 для компьютерной и 398 для низкодозной компьютерной томографии органов грудной клетки. Для окончательного тестирования разработанных алгоритмов были дополнительно размечены тестовые датасеты из 1032 исследований для маммографии, 544 для флюорографии/рентгенографии, 5000 для компьютерной и 1082 для низкодозной компьютерной томографии органов грудной клетки.

Результаты. Наилучшие результаты достигнуты в поиске признаков вирусной пневмонии по протоколам компьютерной томографии органов грудной клетки (точность 0,996, чувствительность 0,998, специфичность 0,989) и рака молочной железы по протоколам маммографии (точность 1,0, чувствительность 1,0, специфичность 1,0). При поиске алгоритмом признаков рака лёгкого метрики получились следующими: точность 0,895, чувствительность 0,829, специфичность 0,936, а при поиске патологических изменений органов грудной клетки в протоколах рентгенографии и флюорографии точность составила 0,912, чувствительность — 1,000, специфичность — 0,844.

Заключение. Машинные методы с высокой точностью могут быть использованы с целью автоматической классификации текстов рентгенологических протоколов маммографии и компьютерной томографии органов грудной клетки для поиска вирусной пневмонии. Для поиска признаков рака лёгкого в модальности компьютерной и низкодозной компьютерной томографии, а также патологических изменений в протоколах рентгенографии и флюорографии органов грудной клетки достигнутой точности достаточно для успешного применения в целях автоматизированного сравнения работы врачей и моделей искусственного интеллекта.

Ключевые слова: протоколы рентгенологических исследований; COVID-19-пневмония; рак лёгкого; рак молочной железы; обработка естественного языка.

Как цитировать

Кокина Д.Ю., Гомболевский В.А., Арзамасов К.М., Андрейченко А.Е., Морозов С.П. Возможности и ограничения использования инструментов машинной обработки текстов в лучевой диагностике // *Digital Diagnostics*. 2022. Т. 3, № 4. С. 374–383. DOI: <https://doi.org/10.17816/DD101099>

Рукопись получена: 18.02.2022

Рукопись одобрена: 24.11.2022

Опубликована: 19.12.2022

DOI: <https://doi.org/10.17816/DD101099>

Possibilities and limitations of using machine text-processing tools in Russian radiology reports

Daria.Yu. Kokina, Victor A. Gombolevskiy, Kirill M. Arzamasov, Anna E. Andreychenko, Sergey P. Morozov

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

ABSTRACT

BACKGROUND: In radiology, important information can be found not only in medical images, but also in the accompanying text descriptions created by radiologists. Identification of study protocols containing certain data and extraction of these data can be useful primarily for clinical problems; however, given the large amount of such data, the development of machine analysis algorithms is necessary.

AIM: To estimate the possibilities and limitations of using a tool for machine processing of radiology reports to search for pathological findings.

MATERIALS AND METHODS: To create an algorithm for automatic analysis of radiology reports, use cases were selected that participated in the experiment on the use of innovative technologies in the computer vision for the analysis of medical images in 2020. Mammography, chest X-ray, chest computed tomography (CT), and LDCT, were among the use cases performed in Moscow. A dictionary of keywords has been compiled. After the automatic marking of the reports by the developed tool, the results were assessed by a radiologist. The number of protocols analyzed by the radiologist for training and validation of the algorithms was 977 for mammography, 4,804 for all chest X-ray scans, 4,074 for chest CT, and 398 for chest LDCT. For the final testing of the developed algorithms, test datasets of 1,032 studies for mammography, 544 for chest X-ray, 5,000 for CT of the chest, and 1,082 studies for the LDCT of the chest were additionally labeled.

RESULTS: The best results were achieved in the search for viral pneumonia in chest CT reports (accuracy 0.996, sensitivity 0.998, and specificity 0.989) and breast cancer in mammography reports (accuracy 1.0, sensitivity 1.0, and specificity 1.0). When searching for signs of lung cancer by the algorithm, the metrics were as follows: accuracy 0.895, sensitivity 0.829, and specificity 0.936, when searching for pathological changes in the chest organs in radiography and fluorography protocols (accuracy 0.912, sensitivity 1.000, and specificity 0.844).

CONCLUSIONS: Machine methods with high accuracy can be used to automatically classify the radiology reports of mammography and chest CT with viral pneumonia. The achieved accuracy is sufficient for successful application to automatically compare the conclusions of physicians and artificial intelligence models when searching for signs of lung cancer in chest CT and LDCT, pathological findings in chest X-ray.

Keywords: radiology reports, COVID-19 pneumonia, lung cancer, breast cancer, natural language processing

To cite this article

Kokina DYu, Gombolevskiy VA, Arzamasov KM, Andreychenko AE, Morozov SP. Possibilities and limitations of using machine text-processing tools in Russian radiology reports. *Digital Diagnostics*. 2022;3(4):374–383. DOI: <https://doi.org/10.17816/DD101099>

Received: 18.02.2022

Accepted: 24.11.2022

Published: 19.12.2022

DOI: <https://doi.org/10.17816/DD101099>

在辐射诊断中使用文本机床的可能性和局限性

Daria.Yu. Kokina, Victor A. Gombolevskiy, Kirill M. Arzamasov, Anna E. Andreychenko, Sergey P. Morozov

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

简评

论证。在放射学中，重要信息不仅包括在医学图像中，还包括在放射科医生创建的随附文本描述中。包含某些数据的研究方案的识别和这些数据的提取首先可能对临床问题有用，但是，鉴于大量此类数据，机器分析算法的开发是必要的。

研究目的是评估使用文本处理工具在放射学协议中搜索病理的可能性和局限性。

材料与方法。为了创建自动协议分析算法的第一个原型，选择了参与使用计算机视觉领域的创新技术进行医学图像分析的实验的研究。这些研究包括在莫斯科医疗机构进行的乳房X光检查、胸部X光摄影、胸部X线间接照相、胸部CT和LDCT。对于每种类型的研究，最初都编制了一个关键词词典，对应于目标病理学的存在与否。在使用开发的工具对协议进行初始自动标记之后，放射科医生对结果进行了选择性评估和验证。医生为训练和验证算法而分析的协议数量为977个乳房X线照相术、3196个射线照相术、1608个荧光照相术、4074个胸部CT和398个胸部LDCT。为了对开发的算法进行最终测试，额外标记了1032项乳房 X线照相术研究、544项荧光照相/射线照相术、5000项胸部CT研究和1082项胸部LDCT研究的测试数据集。

结果。最好结果是根据胸部CT协议（精确度0.996，灵敏度0.998，特异性0.989）和乳房X光检查协议（精确度1.0，灵敏度1.0，特异性1.0）分别在寻找病毒性肺炎迹象和寻找乳腺癌迹象的方面取得的。当通过该算法搜索肺癌征兆时，指标如下：精确度0.895，灵敏度0.829，特异性0.936，以及在射线照相和荧光照相术协议中搜索胸部器官的病理变化时为精确度0.912，灵敏度1.000，特异性0.844。

结论。机器方法可用于乳腺X线检查和胸部CT检查文本的自动分类，以寻找病毒性肺炎。在胸部CT和LDCT模式中寻找肺癌征象，在胸部X线摄影和荧光摄影协议中寻找病理变化，所达到的准确性足以成功应用于医生和人工智能模型工作的自动比较。

关键词：X射线协议；COVID-19肺炎；肺癌；乳腺癌；自然语言处理。

To cite this article

Kokina DYu, Gombolevskiy VA, Arzamasov KM, Andreychenko AE, Morozov SP. 在辐射诊断中使用文本机床的可能性和局限性. *Digital Diagnostics*. 2022;3(4):374–383. DOI: <https://doi.org/10.17816/DD101099>

收到: 18.02.2022

接受: 24.11.2022

发布日期: 19.12.2022

论证

X射线检查报告包含文本医学信息，包括初步诊断、临床数据、具有所研究器官和系统变化特征的描述部分、X射线诊断或结论和观察建议[1, 2]。此信息可用于复杂的疾病诊断和治疗、结果预测和状况监测，以及外部临床目的，例如，用于组织、统计或研究目的。

辐射报告有一系列特点，包括多种叙述风格、电报语音的使用、词汇和术语的变化、多种词序、多种缩写词、首字母缩写词[3]。所有医疗信息的一个共同特点是术语的使用，如果没有经过专业培训，通常无法对术语进行评估。此外，俄语报告还具有一些特定的属性，例如，更宽松的语法规则、词汇多样性；放射科医师使用非标准缩写、复杂的语法结构、难以解释的冗长短语，以及许多否定的变体[4]。词汇变化是放射诊断的特征，然而，在俄语放射诊断中，这种多样性甚至更为广泛（例如，“阴影”可以描述为“黑点”、“影子”、“渗透”、“降低透明度的区域”、“高密度区域”、“通风度降低区域”、“焦点”、“压实”以及针对同一组变化的大量其他变化）。相反，英语放射学的可变性受法规、建议等的约束。因此，放射学检查报告包含大量文本、非结构化和专门的信息，这导致其仅通过自动化方法使用存在一定的困难。

目前，正在研究使用自然语言处理工具来整理报告及使其标准化，突出临床专家所需的信息，自动替换特定术语，包括为患者使用更易理解的词汇或翻译成其他语言[1, 2]。识别包含某些数据的研究报告并提取这些数据主要可用于解决临床问题[1]。例如，有关于识别有骨折迹象的肌肉骨骼描述报告、有肺动脉血栓栓塞迹象的CT扫描、肺部结节等的论文[3, 5, 6]。

为了处理大量的数据，并利用这种分析的结果结合医生的描述和报告来评估医学图像，有必要开发一种机器处理俄语报告的算法。

研究目的是评估使用文本处理工具在放射学协议中搜索各种性质的病理的可能性和局限性。

材料与方 法

放射研究文本报告评估工具的开发

本工作是在先前伦理委员会批准的研究框架内进行的；于2020年2月20日俄罗斯放射学家和X射线学家协会莫斯科地区分会独立伦理委员会（NEC MRO RORR）的第2号记录摘自，同样注册在ClinicalTrials（NCT04489992）。

用于评估放射检查文本协议的工 具是作为莫斯科实验的一部分开发的，该实验使用计算机视觉领域的创新技术来分析医学图像，并用于在人工智能服务和放射科医生的图像上比较评估医学病理学存在的结果。

研究报告包括乳房X线摄影、胸部的X光和荧光摄影（CO）、计算机断层扫描（CT）和低剂量计算机断层扫描（LDCT）；所有的研究都是于2020年在莫斯科卫生部的医疗机构进行的。本研究中使用了匿名放射报告。

主要目标是创建一种自动算法，用于自动分析X射线协议是否存在要探求的病理变化。识别目标病理和开发词典的基础是对人工智能服务结果的基本要求（<https://mosmed.ai/>）。

对于胸部的X光和荧光摄影，目标病理特征包括胸腔积液、气胸、肺不张、病变、浸润/实变、播散、具有腐烂或液位的空腔、钙化和皮质层完整性破坏（骨折）。对于CT和LDCT，目标病理包括体积大于100mm³的实心和亚实心的结节。胸部CT的另一组病变是与冠状病毒感染（COVID-19）体征相关的变化；根据俄罗斯联邦卫生部的临时指南“新型冠状病毒感染（COVID-19）的预防、诊断和治疗”，以及根据国家预算医疗机构莫斯科市卫生局诊断和远程医疗技术科学与实用临床中心的方法学建议“冠状病毒病（COVID-19）X射线诊断：组织、方法、结果解”，使用按严重程度分类[7, 8]。乳腺影像报告和数据系统（BI-RADS 3-6, Breast Imaging Reporting and Data System）被用于乳房X光检查[9]。

文本协议预标记的软硬件解决方案

基于自然语言处理方法并结合放射科医生的专家意见，已经为每种类型的研究开发了文本协议预标记的软件和硬件解决方案。预映射算法的开发是在每个模态的几个主要阶段以迭代方式进行的。

1. 形成一组主要的特征以搜索某些病理学的迹象。所考虑的特征包括关键词和短语、尺寸指示以及必要时的“安全词”和“安全短语”。由放射科医生编制的一套主要关键词，包括放射科医生通用和最常用的术语。安全词包括非目标病理变化或非病理发现（扫描区域其他器官的变化、解剖变异等）。
2. 使用高级编程语言Python将特征集翻译成机器语言。

在本文中，我们使用Python 3.8，需要pandas=1.1.3, numpy=1.19.2, re=2.2.1, nltk=3.5。开发程序的输入数据为表格形式（.csv、.xlsx）的文本报告，其中包含研究描述及其结论。工作结果以标记为“0”或“1”的初始数据呈现，其中“0”表示没有症状/病理，“1”表示存在症状/病理。

用于文本胸部CT扫描的COVID-19特征检索模块是基于自然语言处理（natural language processing, NLP）和机器学习算法系列中的一个分类器。输出表明是否存在COVID-19迹象，以及CT肺损伤程度。

根据乳腺摄影协议寻找乳腺恶性肿瘤体征的模块，包括结论和描述，根据BI-RADS分类检测乳腺恶性肿瘤体征，给出BI-RADS类和二元分

类（类似到筛选量表）作为回应。乳腺学研究结果的描述要求根据BI-RADS量表规定了研究的强制性分类；另一方面，BI-RADS 1的拼写在实践中有多种方案。医生可能会使用不同的大小写、不同的标点符号，最重要的是，不同的布局。由于这个原因，正常的文本解析（parsing）并不总是有效的，从而导致错误的遗漏。NLP在这项任务中的应用归结为信息提取功能。

一个从胸部CT/LDCT文本协议（描述和报告）中搜索肺部恶性肿瘤迹象的模块，从关键词（key）和参数（dimensions）的组合中识别恶性肿瘤的迹象；基于NLP。

在胸部的X光和荧光摄影的报告文本中搜索不同的病理征兆的模块，使用关键词词典（“RG/FLG征兆”：“关键词1”、“关键词2”、“关键词3”...）来识别病理征兆。

3. 使用步骤2中创建的预标记程序标记报告。

4. 在第3阶段从标记样本中选择独特多样（根据病理特征）的程序。在我们的工作过程中，我们一再确信放射科医生使用了死板措词。在死板措词上训练NLP算法将使模型能够快速地被重新训练。由于使用各种稀有（独特）措词丰富了训练数据集，我们已经设定了尽可能广泛地应用所开发算法的任务。

5. 机器标记的人工验证和预标记算法准确性的估计。机器标记的准确性被评估为正确标记的协议的百分比。放射科医生对每个单独的任务进行了多次人工验证。

6. 调整列表的形成，包括额外的“安全词”和短语、关键字和其他建议，以提高预标记的质量。

7. 将经过验证的报告添加到数据库。

在这个阶段，伴有保持类别平衡，形成了多样化、独特和标记的样本。在形成研究样本（样本大小如下）时，使用在有限的时间（从2019年1月到2020年8月）进行的所有研究。各年级的平衡在总体上是一致的。按病理的预期分布：COVID-19 CT：20%正常，80%病理；乳腺X线摄影：95%正常，5%病理；X射线术：95%正常，5%病理；胸透：75%正常，25%病理。

8. 迭代法重复步骤2-7，直到规划针准确度达到98%。这一数值是根据对医学文献进行分析后发现的单个医疗任务的最大NLP精确度（97%）选择的[10]。

放射科医生在制定乳腺造影术前算法时分析报告总数为：乳房造影977个，放射科3196个，透视科1608个，胸部CT 4074个，胸部LDCT 398个。该研究包括所有发送给人工智能服务的研究报告，作为在计算机视觉领域使用创新技术分析医学图像并在莫斯科医疗保健系统中进一步应用的实验的一部分（<https://mosmed.ai/>）。唯一的例外是那些没有完成“描述”和“结论”部分的报告。

为了提高质量并加快文本报告的自动标记，应用机器学习方法考虑报告中复杂的语义句子

结构，从胸部CT扫描中搜索COVID-19肺炎特征。未来，计划开发智能算法以在乳腺X线照相术、CT（用于寻找肺癌）、荧光照相术和射线照相术中寻找病理迹象。

处理COVID-19 CT报告的模块包括三个功能：（1）使用已标记的报告数据库在输入数据中搜索结论；（2）使用常用用语标记其余报告；（3）使用k最近邻（k-Nearest Neighbors, kNN）模型标记其余报告。这些功能的运作得到了一致的执行。我们工具开发的条件之一是快速动作优化。我们以前提到过放射科医生的报告中经常使用死板用语。出于这个原因，大多数COVID-19 CT报告都具有相同的形式。此外，作者团队此前在COVID-19胸部器官靶向病理的CT报告的手动标记方面做了大量的工作。这通过使用一个简单的逻辑比较功能简化和加速了算法，使其能够将协议与数据库中的报告进行比较。然后还有一部分报告不在先前标记研究的数据库中，为此，启动了使用常用词语进行文本分析的慢得多的功能。如果存在正则表达式没有找到目标模式（pattern）的报告，则启动基于kNN的机器学习模型。

与单独的机器学习应用程序相比，我们开发的架构具有速度和准确性的最佳组合。没有机器学习，就不可能涵盖所有报告。学习样本为4074个预先标记的报告。如此多的报告对于确保所需的准确性水平是必要的，并且是在多次测试和重新训练模型的迭代中获得的。为了实现此模块的功能，还导入了Sklearn库（Scikit-learn）。经过训练的算法在测试集上进行了评估，并显示出相当高的准确性（99.6%）。

结果

考虑到报告的特点，已为选定的模式和病理创建了一个关键词和安全词的清单。根据在莫斯科进行的研究方案，在根据胸部CT方案寻找病毒性肺炎COVID-19迹象时，基于开发词典的工具取得了最佳结果：精确度0.996，灵敏度0.998，特异性0.989（TN^{*}=1115；FP^{**}=6；FN[#]=2；TP^{##}=3837）和根据乳房X线照相检查发现的乳腺癌：精确度1.0，敏感性1.0，特异性1.0（TN=461；FP=0；FN=0；TP=571）。在胸部CT和LDCT模式中寻找肺癌征象时，获得以下指标：精确度0.895，敏感性0.829，特异性0.936（TN=619；FP=42；FN=72；TP=349），以及当在胸部的X光和荧光摄影的报告中寻找病理变化分别为0.912；1.000和0.844（TN=259；FP=48；FN=0；TP=237）。

* TN (true negative——真阴性)：预测负类为负类（数），真值为1，预测为1。

** FP (false positive——假阳性)：将负类预测为正类（数），真值为1，预测为0。

FN (false negative——假阴性)：将正类预测为负类（数），真值为0，预测为1。

TP (true positive——真阳性)：预测正类为正类（数），真值为0，预测为0。

讨论

乳腺摄影词典

事实证明，用于乳房X线照相术研究的词典是最容易编写和使用的，这与现代标准化报告的特点有关。描述乳房X光检查的协议是最有条理的，对于这些报告，必须使用BI-RADS，由于检测到的病理的重要性很高，所以它们要受到控制。在绝大多数情况下，BI-RADS类别在结论中指明，而设置该类别或该类别所依据的变化在描述部分。词典中BI-RADS类别的存在使得在报告的自动化处理中实现高精度成为可能，精确度达到1.0。如有必要，该工具可用于其他地区（俄罗斯联邦主体），例如，对于不同结构的报告，可以添加关键定义词（“肿瘤”、“c-r”）。

该工具的操作限制可能与报告中缺少BI-RADS类别有关：通常此类报告包含关于由于图像质量不足或其他原因无法评估乳腺状态的信息。由于此类研究被置于一个单独的组中，因此可以有针对性地进行搜索，例如，用于质量控制，选择需要后续检查的患者。开发的工具，如有进一步改进，可用于其他目的，例如，评估报告与标准的符合性，比较描述性部分和结论以进行审计。

按严重程度划分的COVID-19肺炎研究词汇表

分析COVID-19肺炎严重程度的工具也取得了很高的准确率，这也是由于报告的结构和汇编词汇的明确性。使用的词汇基于标准等级：CT0——没有病毒性肺炎的证据，CT1 - CT4——轻度至重度病毒性肺炎，OTHER——与病毒性肺炎无关的其他变化。在大流行期间，报告包含关于是否存在病毒性肺炎的迹象、与严重程度相关的传播程度以及在大多数情况下发生病毒性肺炎的可能性的信息。在符合规定参数的条件下，该工具只需要报告的文本和基于要评估的损伤程度的拼写变化的词汇。此外，绝大多数报告的描述部分具有共同的特征和共同的术语。

尽管大多数报告具有共同特征，但由于结构、术语的使用、拼写、标点符号、临床医生语言的词汇特点以及个人的CT调查经验的不同，报告的书写形式也有少量的差异。有时会出现与病毒性肺炎CT图像特征相关的诊断错误。在磨玻璃、网状条纹等类型发生变化的情况下，在描述医生看来，它们可能对应于其他疾病，或者存在另一种合并症，协议，特别是报告，可能包含在结构和术语上不典型的句子。在所列举的情况下，该工具的操作可能存在不确定性，使这类研究能够有重点、有针对性地进行审计，找出术语使用的不准确之处，并将其用于临床目的，以确定需要特别注意和需要进一步跟踪的并发症。需要注意的是，在不同的地区和医疗机构中，可以使用不同的分类和报告来描述病毒性肺炎。

肺癌体征查询词典

根据胸部CT和LDCT报告开发关键字和“安全词”来搜索肺癌迹象的任务变得更加困难，导致算法的准确性降低。对LDCT报告有一个Lung-RADS分类，使用该分类可以尽可能地简化寻找可疑结节的任务[11]。然而，在莫斯科进行的研究中搜索Lung-RADS类别并不能对现有的报告进行全面评估，因为并非所有协议的报告都包含这样一个类别；此外，8.3%的报告在描述部分和结论之间存在差异[12]。

开发用于搜索不同性质的肺部结节和肿瘤体征的词典仍然是一项随着许多困难和局限性的紧迫任务。尽管存在用于描述的模板和方法学建议，但胸部器官CT和LDCT的报告具有多种结构和序列选项。许多放射科医生不使用传统的推荐术语（如4毫米的结节被称为“肿块”），这导致了术语的不恰当使用[13]。

考虑到现有的建议和期待的术语，采用关键字来搜索可疑的变化，对应于6毫米的肺实性结节/病灶，以及结构（大小变化超过3厘米）[10-12]。这些标准有一些局限性，会导致算法出现假阳性或假阴性。例如，在进行胸部CT扫描时，建议使用费莱舍尔学会（Fleischner）的建议来评估意外发现的实体肺结节，但是，其应用需要评估护照数据、临床信息、评估风险因素，是否存在其他疾病，包括肿瘤特征[14]。

在使用规模标准和一组核心关键词时，不能完全排除良性变化。例如，要排除结构中钙化的良性节点，必须考虑其分布的性质，这通常在报告中未指定。同时，描述钙化可以是在恶性肿瘤的结构中描述的。医生可以将大病灶描述为其他疾病的一部分，例如，各种病因的结核病、结节病、细支气管炎。

所开发算法的现有能力足以完成比较报告数据和人工智能服务的处理结果的任务。进一步改进该工具，提高准确性，不仅可以将其用于其他地区，还有助于为其他任务创建有用的工具。

随着工具的改进，可以创建一个更准确的算法，考虑到必要的风险因素（免疫缺陷的存在、报告中的炎症过程、临床信息、转诊诊断等）。这种改善在评估癌症患者的肺结节时可能尤为重要。关于是否有癌症的信息可以从报告的描述部分以及电子病历中获得的。使用该工具估计结节大小变化的动态，并将肺结节管理建议与数据进行比较，这也可能是改进该工具以使其功能与改进计算机视觉模型相匹配的一个有前途的方向。

一组单独的限制，尤其与胸部CT组相关，与广泛且难以涵盖的“安全词”列表相关。这是由于这项研究的特殊性：腹腔、颈部以及胸部其他器官都落入了扫描区域。对于大多数器官（例如，甲状腺、肝脏、肾脏、肾上腺等），此类“安全词”包括本身器官的名称。然而，胸腔解剖结构，如纵隔、心包、肋骨和胸椎、胸壁软组织和横膈膜不能作为独立的“安全词”，因为肺部恶

性肿瘤有侵袭性生长并影响邻近组织的情况并不罕见，医生将其进行总体描述（“肿块扩散至纵隔”）。同时，经常遇到所列器官和解剖结构的各种独立肿瘤，导致算法出现大量误报。

胸部X射线和荧光摄影关键词表

开发和应用穿刺文本分析工具的一个具有挑战性的领域是胸部放射摄影和荧光摄影。X射线和荧光摄影协议在形状、结构、大小、使用特性方面有许多变化，而术语也有很大差异[15, 16]。

制定的胸部X光摄影和透视的关键词表，除了常见的病理标签以外，还包括具体的放射学术语，如“黑点”、“焦点”和“阴影”。这与许多困难相关，因为此类术语可用于非目标病理学或解剖结构（“肋骨阴影”）、其他医疗设备（“起搏器阴影”、“引流管阴影”）的符号。

根据具有所列困难情况的二元分类（正常/病理）的病理学定义可能具有相当高的准确性，但是，如果有必要将病理学分为许多类别（例如，渗出物、气胸、肺不张、病灶等），即使是放射科医师自己也很难将结果分类，这也与X射线方法的能力有限有关。此外，相同的关键字用于指代完全不同的放射学发现。

在开发“安全词”的过程中，考虑了使用各种否定的词汇和句法变体的多种规范拼写（“肺部无黑点”、“肺部无病理阴影”等）。随着新数据的到来，需要不断扩充词典。目前，为此类研

究获得的准确性指标使我们能够解决比较人工智能模型和医生结果的主要问题。

结论

机器方法具有很高的准确性，可用于自动分类乳房X线摄影和胸部CT的X射线报告文本以搜索病毒性肺炎，这与这些研究的描述的结构化和标准化有关。

在胸部CT和LDCT研究协议中寻找肺癌征象，在胸部器官的射线照相和荧光造影协议中寻找病理变化，所达到的准确性足以成功使用，以便自动比较医生的工作和放射科的人工智能模型。较低的准确性是由于协议的更自由的结构、其诊断、词汇和术语的特征。

ADDITIONAL INFORMATION

Funding source. This study was not supported by any external sources of funding.

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.

СПИСОК ЛИТЕРАТУРЫ

1. Sorin V., Barash Y., Konen E., Klang E. Deep learning for natural language processing in radiology: Fundamentals and a systematic review // *J Am College Radiol.* 2020. Vol. 17, N 5. P. 639–648. doi: 10.1016/j.jacr.2019.12.026
2. Monshi M.M., Poon J., Chung V. Deep learning in generating radiology reports: A survey // *Artif Intell Med.* 2020. N 106. P. 101878. doi: 10.1016/j.artmed.2020.101878
3. Banerjee I., Chen M.C., Lungren M.P., Rubin D.L. Radiology report annotation using intelligent word embeddings: Applied to multi-institutional chest CT cohort // *J Biomed Inform.* 2018. N 77. P. 11–20. doi: 10.1016/j.jbi.2017.11.012
4. Kivotova E., Maksudov B., Kuleev R., Ibragimov B. Extracting clinical information from chest x-ray reports: A case study for Russian language // *Conference: International Conference Nonlinearity, Information and Robotics (NIR)At: Innopolis, Russia, 2020.* P. 1–6. doi: 10.1109/NIR50484.2020.9290235
5. Lee C., Kim Y., Kim Y.S., Jang J. Automatic disease annotation from radiology reports using artificial intelligence implemented by a recurrent neural network // *Am J Roentgenol.* 2019. Vol. 212, N 4. P. 734–740. doi: 10.2214/AJR.18.19869
6. Yuan J., Zhu H., Tahmasebi A. Classification of pulmonary nodular findings based on characterization of change using radiology reports // *AMIA Jt Summits Transl Sci Proc.* 2019. Vol. 2019. P. 285–294.
7. 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. Москва, 2020. 78 с.

8. Профилактика, диагностика и лечение новой коронавирусной инфекции 2019-нCoV. Временные методические рекомендации Министерства здравоохранения Российской Федерации // *Пульмонология.* 2019. Т. 29, № 6. С. 655–672. doi: 10.18093/0869-0189-2019-29-6-655-672

9. D'Orsi C.J., Sickles E.A., Mendelson E.B., et al. ACR BI-RADS atlas, breast imaging reporting and data system. Reston, VA, American College of Radiology; 2013.

10. Caliskan D., Zierk J., Kraska D., et al. First steps to evaluate an NLP tool's medication extraction accuracy from discharge letters // *Stud Health Technol Inform.* 2021. N 278. P. 224–230. doi: 10.3233/SHTI210073

11. American College of Radiology Committee on Lung-RADS. Lung-RADS Assessment Categories version 1.1. Режим доступа: <https://www.acr.org/-/media/ACR/Files/RADS/Lung-RADS/LungRADSAssessmentCategoriesv1-1.pdf>. Дата обращения: 01.01.2020.

12. Морозов С.П., Владимирский А.В., Гомболевский В.А., и др. Искусственный интеллект: автоматизированный анализ текста на естественном языке для аудита радиологических исследований // *Вестник рентгенологии и радиологии.* 2018. Т. 99, № 5. С. 253–258.

13. Hansell D.M., Bankier A.A., MacMahon H., et al. Fleischner Society: glossary of terms for thoracic imaging // *Radiology.* 2008. Vol. 246, N 3. P. 697–722. doi: 10.1148/radiol.2462070712

14. MacMahon H., Naidich D.P., Goo J.M., et al. Guidelines for management of incidental pulmonary nodules detected on CT images: from the Fleischner Society 2017 // *Radiology*. 2017. Vol. 284, N 1. P. 228–243. doi: 10.1148/radiol.2017161659

15. Callister M.E., Baldwin D.R., Akram A.R., et al. British Thoracic Society Pulmonary Nodule Guideline Development Group; British Thoracic Society Standards of Care Committee. British Tho-

racic Society guidelines for the investigation and management of pulmonary nodules // *Thorax*. 2015. Vol. 70, Suppl 2. P. ii1–ii54. doi: 10.1136/thoraxjnl-2015-207168

16. Синицын В.Е., Комарова М.А., Мершина Е.А. Протокол рентгенологического описания: прошлое, настоящее, будущее // *Вестник рентгенологии и радиологии*. 2014. № 3. С. 35–40.

REFERENCES

1. Sorin V, Barash Y, Konen E, Klang E. Deep learning for natural language processing in radiology: Fundamentals and a systematic review. *J Am Coll Radiol*. 2020;17(5):639–648. doi: 10.1016/j.jacr.2019.12.026

2. Monshi MM, Poon J, Chung V. Deep learning in generating radiology reports: A survey. *Artif Intell Med*. 2020;(106):101878. doi: 10.1016/j.artmed.2020.101878

3. Banerjee I, Chen MC, Lungren MP, Rubin DL. Radiology report annotation using intelligent word embeddings: Applied to multi-institutional chest CT cohort. *J Biomed Inform*. 2018;(77):11–20. doi: 10.1016/j.jbi.2017.11.012

4. Kivotova E, Maksudov B, Kulee R, Ibragimov B. Extracting clinical information from chest X-ray reports: A case study for Russian language. Conference: International Conference Nonlinearity, Information and Robotics (NIR)At: Innopolis, Russia; 2020. P. 1–6. doi: 10.1109/NIR50484.2020.9290235

5. Lee C, Kim Y, Kim YS, Jang J. Automatic disease annotation from radiology reports using artificial intelligence implemented by a recurrent neural network. *Am J Roentgenol*. 2019;212(4):734–740. doi: 10.2214/AJR.18.19869

6. Yuan J, Zhu H, Tahmasebi A. Classification of pulmonary nodular findings based on characterization of change using radiology reports. *AMIA Jt Summits Transl Sci Proc*. 2019;2019:285–294.

7. Morozov SP, Protsenko DN, Smetanina SV, et al. Radiation diagnostics of coronavirus disease (COVID-19): Organization, methodology, interpretation of results: Preprint, 2020-II. Version 2. Moscow; 2020. 78 p. (In Russ).

8. The prevention, diagnosis and treatment of the new coronavirus infection 2019-nCoV. Temporary guidelines Ministry of Health of the Russian Federation. *Pulmonologiya*. 2019;29(6):655–672. (In Russ). doi: 10.18093/0869-0189-2019-29-6-655-672

9. D'Orsi CJ, Sickles EA, Mendelson EB, et al. ACR BI-RADS Atlas, Breast Imaging Reporting and Data System. Reston, VA, American College of Radiology; 2013.

10. Caliskan D, Zierk J, Kraska D, et al. First steps to evaluate an NLP tool's medication extraction accuracy from discharge letters. *Stud Health Technol Inform*. 2021;(278):224–230. doi: 10.3233/SHTI210073

11. American College of Radiology Committee on Lung-RADS. Lung-RADS Assessment Categories version 1.1. Available from: <https://www.acr.org/-/media/ACR/Files/RADS/Lung-RADS/LungRADSAssessmentCategoriesv1-1.pdf>. Accessed: 01.01.2020.

12. Morozov SP, Vladimirovskiy AV, Gombolevskiy VA, et al. Artificial intelligence: natural language processing for peer-review in radiology. *J Radiol Nuclear Med*. 2018;99(5):253–258. (In Russ). doi: 10.20862/0042-4676-2018-99-5-253-258

13. Hansell DM, Bankier AA, MacMahon H, et al. Fleischner Society: glossary of terms for thoracic imaging. *Radiology*. 2008;246(3):697–722. doi: 10.1148/radiol.2462070712

14. MacMahon H, Naidich DP, Goo JM, et al. Guidelines for management of incidental pulmonary nodules detected on CT images: From the Fleischner Society 2017. *Radiology*. 2017;284(1):228–243. doi: 10.1148/radiol.2017161659

15. Callister ME, Baldwin DR, Akram AR, et al.; British Thoracic Society Pulmonary Nodule Guideline Development Group; British Thoracic Society Standards of Care Committee. British Thoracic Society guidelines for the investigation and management of pulmonary nodules. *Thorax*. 2015;70(Suppl 2):ii1–ii54. doi: 10.1136/thoraxjnl-2015-207168

16. Sinityn VE, Komarova MA, Mershina EA. Radiology report: past, present and future. *J radiol nuclear med*. 2014;(3):35–40. (In Russ).

AUTHORS' INFO

* Daria Yu. Kokina;

address: 24-1 Petrovka str., 127051, Moscow, Russia;
ORCID: <https://orcid.org/0000-0002-1141-8395>;
eLibrary SPIN: 9883-4656; e-mail: d.kokina@npcmr.ru

Victor A. Gombolevskiy, MD, Cand. Sci. (Med.);

ORCID: <https://orcid.org/0000-0003-1816-1315>;
eLibrary SPIN: 6810-3279; e-mail: g_victor@mail.ru

ОБ АВТОРАХ

* Кокина Дарья Юрьевна;

адрес: Россия, 127051, Москва, ул. Петровка, д. 24, стр. 1;
ORCID: <https://orcid.org/0000-0002-1141-8395>;
eLibrary SPIN: 9883-4656; e-mail: d.kokina@npcmr.ru

Гомболевский Виктор Александрович, к.м.н.;

ORCID: <https://orcid.org/0000-0003-1816-1315>;
eLibrary SPIN: 6810-3279; e-mail: g_victor@mail.ru

* Corresponding author / Автор, ответственный за переписку

Kirill M. Arzamasov, MD, Cand. Sci. (Med.);
ORCID: <http://orcid.org/0000-0001-7786-0349>;
eLibrary SPIN: 3160-8062; e-mail: k.arzamasov@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

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

Арзамасов Кирилл Михайлович, к.м.н.;
ORCID: <http://orcid.org/0000-0001-7786-0349>;
eLibrary SPIN: 3160-8062; e-mail: k.arzamasov@npcmr.ru

Андрейченко Анна Евгеньевна, канд. физ.-мат. наук;
ORCID: <https://orcid.org/0000-0001-6359-0763>;
eLibrary SPIN: 6625-4186; e-mail: a.andreychenko@npcmr.ru

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