

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ: АВТОМАТИЗИРОВАННЫЙ АНАЛИЗ ТЕКСТА НА ЕСТЕСТВЕННОМ ЯЗЫКЕ ДЛЯ АУДИТА РАДИОЛОГИЧЕСКИХ ИССЛЕДОВАНИЙ

С.П. Морозов, доктор мед. наук, профессор, директор

А.В. Владзимирский, доктор мед. наук, заместитель директора по научной работе

В.А. Гомболевский, канд. мед. наук, руководитель отдела развития качества радиологии

Е.С. Кузьмина, заместитель директора по общим вопросам

Н.В. Ледихова, заведующая консультативным отделением

ГБУЗ г. Москвы «Научно-практический центр медицинской радиологии Департамента здравоохранения города Москвы»,
ул. Средняя Калитниковская, 28, стр. 1, Москва, 109029, Российская Федерация

ARTIFICIAL INTELLIGENCE: NATURAL LANGUAGE PROCESSING FOR PEER-REVIEW IN RADIOLOGY

S.P. Morozov, Dr. Med. Sc., Professor, Director; orcid.org/0000-0001-6545-6170

A.V. Vladzimirskiy, Dr. Med. Sc., Deputy Director for Scientific Work; orcid.org/0000-0002-2990-7736

V.A. Gombolevskiy, Cand. Med. Sc., Head of Radiology Quality Development Department;
orcid.org/0000-0003-1816-1315

E.S. Kuz'mina, Deputy Director for General Issues; orcid.org/0000-0003-0235-9386

N.V. Ledikhova, Head of Consultative Department; orcid.org/0000-0002-1446-424X

Research and Practical Center of Medical Radiology, Department of Health of Moscow,
ul. Srednyaya Kalitnikovskaya, 28, stroenie 1, Moscow, 109029, Russian Federation

Цель исследования – оценить значимость системы обработки естественного языка для анализа качества протоколов радиологических исследований.

Материал и методы. На базе коммерчески доступной когнитивной системы обработки естественного языка проведен многосторонний анализ протоколов низкодозных компьютерных томографий (НДКТ) органов грудной клетки. Выполнена оценка применимости искусственного интеллекта для выявления расхождений в описаниях исследований (количественный анализ) и для оценки приверженности врачей-радиологов рекомендациям по ведению очагов в соответствии с Lung-RADS-2014 (качественный анализ).

Результаты. Согласно результатам количественного анализа, в 8,3% протоколов НДКТ содержались расхождения между описанием и заключением. Суть расхождений – значимый элемент, например наличие очагов в легких, указан лишь в одном компоненте протокола. Данное расхождение несет потенциальные риски и должно учитываться в процессе аудита качества радиологических исследований. Результаты качественного анализа: для очагов Lung-RADS 3 рекомендованные принципы ведения пациентов использованы в 46% случаев, для Lung-RADS 4A – в 42%, а для Lung-RADS 4B – в 49%.

Заключение. Согласованность решений при использовании системы обработки естественного языка в рамках аудита радиологических исследований составляет 95–96%. Можно констатировать факт применимости системы обработки естественного языка в качестве инструмента для аудита радиологических исследований.

Ключевые слова: радиология; низкодозная компьютерная томография; искусственный интеллект; контроль качества; обработка естественного языка.

Для цитирования: Морозов С.П., Владзимирский А.В., Гомболевский В.А., Кузьмина Е.С., Ледихова Н.В. Искусственный интеллект: автоматизированный анализ текста на естественном

Objective. To assess the importance of natural language processing (NLP) system for quality assurance of the radiological reports.

Material and methods. Multilateral analysis of chest low-dose computed tomography (LDCT) reports based on a commercially available cognitive NLP system was performed. The applicability of artificial intelligence for discrepancy identification in the report body and conclusion (quantitative analysis) and radiologist adherence to the Lung-RADS guidelines (qualitative analysis) was evaluated.

Results. Quantitative analysis: in the 8.3% of cases LDCT reports contained discrepancies between text body and conclusion, i.e., lung nodule described only in body or conclusion. It carries potential risks and should be taken into account when performing a radiological study audit. Qualitative analysis: for the Lung-RADS 3 nodules, the recommended principles of patient management were used in 46%, for Lung-RADS 4A – in 42%, and for Lung-RADS 4B – in 49% of cases.

Conclusion. The consistency of NLP system within the framework of radiological study audit was 95–96%. The system is applicable for the radiological study audit, i.e. large-scale automated analysis of radiological reports and other medical documents.

Index terms: radiology; low-dose computed tomography; artificial intelligence; quality control; natural language processing.

For citation: Morozov S.P., Vladzimirskiy A.V., Gombolevskiy V.A., Kuz'mina E.S., Ledikhova N.V. Artificial intelligence: natural language processing for peer-review in radiology. *Journal of Radiology and Nuclear Medicine*. 2018; 99 (5): 253–8 (in Russ.). DOI: 10.20862/0042-4676-2018-99-5-253-258

For correspondence: Anton V. Vladzimirskiy; E-mail: a.vladzimirskiy@nrcpmf.ru

Введение

Интенсивность использования компьютерной и магнитно-резонансной томографии (КТ, МРТ) нарастает стремительно, такая тенденция обуславливает необходимость системного контроля качества и безопасности выполняемых исследований. Концепция аудита результатов радиологических исследований начала формироваться несколько лет назад. Пересмотр результатов обследований применяется для контроля качества службы лучевой диагностики, а также аттестации медицинских работников. В мире аудит стал стандартным инструментом управления качеством на уровне медицинской организации [1–3].

Принципиально более прогрессивная стратегия была реализована ГБУЗ «Научно-практический центр медицинской радиологии Департамента здравоохранения г. Москвы» (далее НППЦ МР). В 2016–2017 гг. был создан Единый радиологический информационный сервис (ЕРИС) – информационная система в сфере здравоохранения, объединяющая рабочие места специалистов и диагностическую аппаратуру медицинских организаций (МО). ЕРИС позволяет мониторировать работу службы лучевой диагностики в режиме реального времени. Штатными врачами-экспертами НППЦ МР проводится постоянный телеаудит результатов радиологических исследований; выборка формируется автоматически, случайным образом, ее объем составляет до 10% от общего числа исследований по обеим модальностям (КТ и МРТ), выполненным в данной МО за определенный период времени. Результативность и значимость этой деятельности отражена в ра-

нее опубликованных наших материалах [4–6].

Большой объем данных (за 2017 г. в ЕРИС накоплено около 1 млн исследований) буквально вынуждает работать над решением проблемы эффективной автоматизации анализа диагностических изображений и медицинских документов. Разработка алгоритмов, инструментов распознавания и анализа преследует сугубо практическую задачу – обеспечить аудит 100% радиологических исследований, что позволит перевести всю службу лучевой диагностики столицы на принципиально новый уровень качества. Результатом КТ или МРТ является набор изображений в формате DICOM и текстовый документ – протокол, включающий в том числе описание и заключение. В связи с этим автоматический анализ данных разделяется на два направления – распознавание изображений и обработку естественного языка (которые входят в концепцию искусственного интеллекта). И если первое направление достаточно широко изучается в медицине, то значение второго явно недооценено, во всяком случае, для аудита в радиологии.

В последние годы на мировой рынок вышли новые системы обработки естественного языка от флагманов индустрии информационных технологий [7]. Такие инструменты уже успешно применяются в науке, бизнесе и даже в социальной сфере. Чрезвычайная актуальность данного направления и определила цель нашего исследования – оценить значимость системы обработки естественного языка для анализа качества протоколов радиологических исследований.

Материал и методы

Исследование выполнено на базе НППЦ МР. Было поставлено две задачи: оценить применимость системы для выявления расхождений в описаниях исследований (количественный анализ), оценить приверженность врачей-радиологов рекомендациям по ведению очагов в соответствии с классификацией Lung-RADS в версии 2014 г. (качественный анализ соответствия рекомендаций выявленным типам очагов) [8].

Использованы результаты 5047 низкодозных компьютерных томографий (НДКТ) легких, выполненных в 2017 г. в государственных городских поликлиниках Департамента здравоохранения г. Москвы и помещенных в ЕРИС [9, 10]. Выбор именно результатов НДКТ обусловлен тем, что данный метод применяется для массового скрининга рака легких, а значит, требует первоочередной интеграции средств автоматического анализа данных [9]. Врачи-рентгенологи при формировании текстов протоколов НДКТ брали за основу шаблон без патологических изменений [11]. Сформировано три выборки, включавших протоколы НДКТ (каждый протокол состоял из описания и заключения). Все документы были деперсонализированы в соответствии с требованиями законодательства Российской Федерации, а также лишены информации о медицинском учреждении и технических характеристик выполненного исследования. Выборка для обучения и валидации системы автоматического анализа состояла из 395 протоколов, выборка для количественного анализа – из 3052, а для качественного – 5047 протоколов. Раз-

деление проводилось случайным образом.

На базе коммерчески доступной когнитивной системы обработки естественного языка проведен многосторонний анализ протоколов НДКТ органов грудной клетки с целью обнаружения в них изменений по заданным параметрам, а также для выявления неточностей.

Обучение и настройка системы осуществлялась сотрудниками компании-разработчика. Специалисты НПЦ МР обеспечили формирование выборок, подготовку необходимых словарей и классификаторов, определение критично значимых слов и их сочетаний (триггеров), проведение консультаций в процессе обучения системы и оценки точности. В процессе многоэтапной работы в библиотеки системы были внесены уточнения, повысившие точность выявления триггеров.

Оценка согласованности классификаций выполнена путем расчета коэффициента каппа Коэна.

Результаты

Оценка применимости алгоритмов обработки естественного языка для выявления расхождений в описаниях исследований (количественный анализ). Для количественного анализа были поставлены следующие задачи: охарактеризовать выявляемость триггеров, выявить расхождения в наличии триггеров между компонентами протокола (описанием и заключением), оценить согласованность выявления расхождений при автоматизированном анализе и при независимом анализе квалифицированным радиологом.

Для тестирования результатов работы прототипа был проведен анализ 3052 деперсонализированных протоколов НДКТ. В результате в 457 (14,9%) протоколах автоматически обнаружены 1412 триггеров, связанных с легкими: наиболее часто – «очаг» (51,0%) и «уплотнение» (24,0%), наиболее редко – «фо-

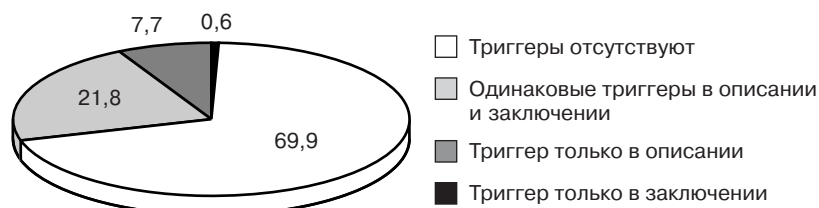


Рис. 1. Удельный вес триггеров в компонентах заключения (по результатам анализа протоколов низкодозных компьютерных томографий системой обработки естественного языка)

Сводные результаты анализа НДКТ протоколов скрининга рака легкого системой обработки естественного языка

Тип сущности	Количество уникальных значений	Количество протоколов, в которых выявлены сущности
Потенциальные очаги	7	4799 (95,08%)
Новый контекст	72	5027 (99,06%)
Очаги в легких	7	1690 (33,48%)
Свойства очагов	429	1690 (33,48%)
Рекомендации	43	3936 (77,99%)
Нерелевантные очаги	33	4144 (82,11%)
Lung-RADS: классы 3, 4A, 4B	3	350 (6,93%)
Выявление несоответствий описания и заключения	4	4830 (95,70%)
Общее количество протоколов НДКТ	–	5047 (100%)

кус» (0,6%) и «диссеминированный» (0,2%). Количество обнаруженных триггеров составило: один – в 378 протоколах, два – в 61, более двух – в 17, множество – в 1. Среди триггеров, характеризующих контуры очагов, преобладали «четкий», «ровный», среди характеристик формы – «округлый», «неправильный». Триггеры, относящиеся к типу очага, распределились таким образом: «солидный» – 56,0%, «полусолидный» – 23,0%, «матовое стекло» – 20,0%, «субсолидный» – 1,0%.

Результаты автоматизированного анализа наличия триггеров в компонентах протоколов НДКТ представлены на диаграмме (рис. 1).

В результате автоматизированного анализа протоколов НДКТ установлено, что в 8,3% документов содержались расхождения между описанием и заключением. Суть расхождений состояла в том, что значимый элемент (например, наличие очагов в легких) был указан лишь

в одном компоненте протокола. Данное расхождение несет потенциальные риски и должно учитываться в процессе аудита качества радиологических исследований.

При математическом сопоставлении согласованности решений о расхождениях в наличии триггеров между компонентами протокола установлено, что коэффициент каппа Коэна составляет 0,95. Следовательно, согласованность решений системы обработки естественного языка и квалифицированного врача-радиолога может быть оценена как высокая.

Оценка приверженности врачей-радиологов рекомендациям по ведению очагов в соответствии с классификацией Lung-RADS (качественный анализ). Определено соответствие рекомендаций, содержащихся в заключении, типу выявленного очага по классификации Lung-RADS [12] (см. таблицу).

На первом этапе проанализировано 5047 протоколов НДКТ,

автоматически выявлено 3936 документов с различными рекомендациями, в том числе 1690 – с описаниями очагов. Из этого числа 100 (5,9%) классифицировано как содержащие описание очагов Lung-RADS 3, 161 (9,5%) – как Lung-RADS 4A и 89 (5,3%) – как Lung-RADS 4B.

На втором этапе были проанализированы рекомендации относительно типов очагов (рис. 2).

Установлено, что для очагов Lung-RADS 3 рекомендованные принципы ведения пациентов использованы в 46%, для Lung-RADS 4A – в 42%, а для Lung-RADS 4B – в 49% случаев.

Автоматическая классификация очагов и анализ заключений были верифицированы двумя независимыми врачами-радиологами; точность системы обработки естественного языка составила 96%.

Обсуждение

Алгоритмы обработки естественного языка достаточно давно применяются в радиологии для получения структурированных данных из протоколов, подготовленных врачами, в рамках отдельных научных исследований [12–17]. При этом реальное внедрение алгоритмов в повседневную клиническую работу встречается редко (примерно 30% случаев) [18]. В 2017 г. показана достаточная ценность алгоритмов обработки естественного языка для анализа протоколов радиологических исследований с целью выявления признаков тромбозов [19, 20]. В том же году разработан и апробирован алгоритм анализа клинических данных для автоматического определения показаний к выполнению МРТ с внутривенным контрастированием при травмах и заболеваниях опорно-двигательной системы. Точность работы алгоритма составила 83,2–90,0%, чувствительность – 0,743–0,812, специфичность – 0,921–0,952. Исследователи отметили особую надежность алгоритмов обработ-

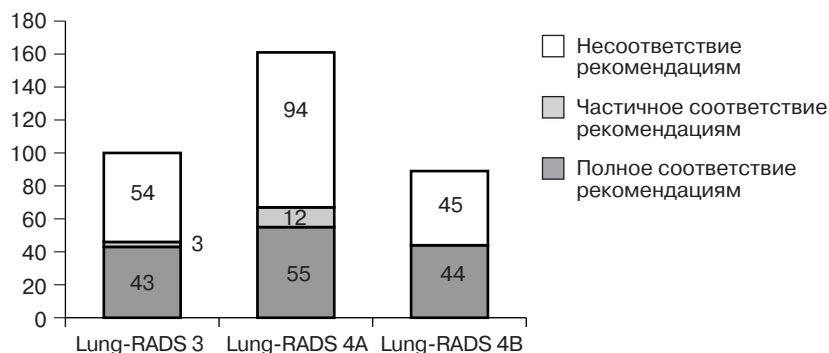


Рис. 2. Приверженность врачей-радиологов рекомендациям по ведению очагов в соответствии с классификацией Lung-RADS (по результатам анализа протоколов низкодозных компьютерных томографий системой обработки естественного языка)

ки естественного языка в отсеивании грамматических и орфографических ошибок, часто встречающихся в протоколах. Был сделан вывод о возможности создания клинической системы поддержки принятия решений, потенциально позволяющей снизить количество ошибок и оптимизировать работу как радиологов, так и врачей-клиницистов [21]. Показана возможность автоматизации аудита, проводимого в отдельно взятом отделении лучевой диагностики, посредством системы обработки естественного языка. На ограниченной выборке из 25 протоколов точность определения использования врачом структурированного описания составила 92,0% при машинном анализе, 91,2% – при анализе врачами-экспертами [22].

Неизученным остается вопрос применения систем обработки естественного языка в процессе проведения массовых профилактических осмотров, например скрининга рака легкого посредством НДКТ [9, 23, 24]. В связи с этим можно обратить внимание лишь на публикацию S.E. Beyer et al. (2017 г.). Авторы разработали алгоритмы, позволяющие извлекать необходимые данные из структурированных описаний результатов НДКТ и формировать на их основе рекомендации по классификации очагов (в соответствии с Lung-RADS). Точность системы достигает 87%, специфичность – 99,5%

[25]. В данном случае речь идет о формировании системы поддержки принятия решений, но не о контроле качества.

Тема автоматизированного анализа медицинской документации в лучевой диагностике является очень актуальной. При этом лидируют исследования по созданию систем поддержки принятия решений. В то же время тема контроля качества фактически не изучена; единичные работы посвящены анализу ограниченных выборок в контексте сугубо внутреннего аудита отделений лучевой диагностики [18, 22].

В нашем исследовании алгоритмы анализа естественного языка применены для проведения централизованного контроля качества в сетях медицинских организаций, в том числе при массовых профилактических осмотрах и использовании неструктурированных описаний.

На первом этапе была изучена выявляемость триггеров – расхождений между компонентами протокола – посредством системы искусственного интеллекта. Установлено, что в 8,3% документов содержались расхождения между описанием и заключением (согласованность решений составляет 95%).

Полученные результаты по согласованности в целом соответствуют литературным данным [21, 22]. Подтверждены выводы об эффективности искусственного интеллекта в устранении тех-

нических ошибок при оформлении протоколов [21].

На втором этапе нами изучена приверженность врачей-радиологов рекомендациям по ведению очагов в соответствии с Lung-RADS. Как было отмечено выше, для очагов Lung-RADS 3 рекомендованные принципы ведения пациентов использованы в 46% случаев, для Lung-RADS 4A – в 42%, а для Lung-RADS 4B – в 49% (точность системы обработки естественного языка 96%). Полученные данные превосходят результаты аналогичного исследования, посвященного анализу приверженности радиологов рекомендациям Флейшнеровского общества при выявлении очагов в легких [26]. Посредством системы обработки естественного языка проанализировано 1100 КТ, в результате выявлено 314 протоколов, содержащих описание очагов. При качественном анализе установлено, что рекомендованные принципы ведения пациентов использованы только в 34% протоколов. Среди причин такой ситуации авторы указали неофициальный статус рекомендаций, отсутствие активной поддержки принятия решений и информирования врачей. При этом приверженность рекомендациям статистически зависела от размера выявленного очага. При наличии очагов размером 4–6 мм зафиксировано достоверно более частое следование методическим принципам Флейшнеровского общества.

Таким образом, качество выявления системой обработки естественного языка зависимостей между типом очага по классификации Lung-RADS и рекомендациями врача можно оценить как высокое. Полученные нами данные о характере следования методическим рекомендациям врачами, участвующими в популяционном скрининге рака легкого, превосходят показатели зарубежных авторов.

Системы обработки естественного языка могут применяться

для централизованного аудита неструктурированных протоколов радиологических исследований, проводимых в сетях медицинских организаций, в том числе в рамках массовых профилактических осмотров.

Дальнейшие исследования могут быть направлены на применение алгоритмов анализа естественного языка для оценки 100% исследований при скрининге рака легкого. Представленные результаты являются одним из шагов к развитию более сложных когнитивных систем, объединяющих использование алгоритмов обработки естественного языка, структурированный подход к имеющейся информации и результатам анализа изображений лучевой диагностики для применения в реальном времени при первичном формировании протокола исследования.

Ограничения исследования.

В данной работе не проводился сравнительный анализ различных систем обработки естественного языка, не рассматривались юридические вопросы автоматизированного контроля качества в медицинских организациях, а также не выполнялся анализ характеристической кривой. Все эти вопросы будут изучены в дальнейших исследованиях.

Заключение

Контроль качества радиологических исследований – актуальная проблема современного здравоохранения, эффективно решаемая только путем применения телемедицинских и иных информационных технологий, ориентированных прежде всего на автоматизированный анализ медицинских данных.

Алгоритмы коммерческой системы обработки естественного языка позволили провести автоматизированный анализ 5047 депersonализированных протоколов низкодозных компьютерных томографий, выполненных в рамках масштабной программы скрининга рака легких. согласо-

ванность решений при использовании системы в рамках аудита радиологических исследований составила 95–96%.

Установлены потенциальные возможности системы для выявления значимых слов и их сочетаний, характеризующих качество описаний и заключений, а также для анализа соответствия рекомендаций, формулируемых врачом-радиологом, методическим и иным требованиям.

Можно констатировать факт применимости системы обработки естественного языка в качестве инструмента для аудита радиологических исследований, то есть для сплошного автоматизированного анализа протоколов и иных текстовых документов.

Конфликт интересов

Авторы заявляют об отсутствии конфликта интересов.

Литература [References]

1. Brady A. Error and discrepancy in radiology: inevitable or avoidable? *Ins. Imag.* 2017; 1 (8): 171–82.
2. Larson D.B., Donnelly L.F., Podberesky D.J., Merrow A.C., Sharpe R.E. Jr, Kruskal J.B. Peer feedback, learning, and improvement: answering the call of the institute of medicine report on diagnostic error 1. *Radiology.* 2017; 283 (1): 231–41. DOI: 10.1148/radiol.2016161254
3. Walker E.A., Petscavage-Thomas J.M., Fotos J.S., Bruno M.A. Quality metrics currently used in academic radiology departments: results of the QUALMET survey. *Br. J. Radiol.* 2017; 90 (1071): 20160827. DOI: 10.1259/bjr.20160827
4. Morozov S., Guseva E., Ledikhova N., Vladzimirsky A., Safonov D. Telemedicine-based system for quality management and peer review in radiology. *Ins. Imag.* 2018; 9 (3): 337–41. DOI: 10.1007/s13244-018-0629-y
5. Морозов С.П., Переверзев М.О. Лучевая диагностика – авангард информатизации здравоохранения. *Российский электронный журнал лучевой диагностики.* 2013; 3: 41–50. [Morozov S.P., Pereverzev M.O. Radiology – avanguard of healthcare informatization. *Rossiyskiy Elektronnyy Zhurnal Luchevoy Diagnostiki (Russian Electronic Journal of Radiology).* 2013; 3: 41–50 (in Russ.).]

6. Полищук Н.С., Ветшева Н.Н., Косарин С.П., Морозов С.П., Кузьмина Е.С. Единый радиологический информационный сервис как инструмент организационно-методической работы Научно-практического центра медицинской радиологии Департамента здравоохранения г. Москвы. *Радиология–практика*. 2018; 1 (67): 6–17. [Polishchuk N.S., Vetsheva N.N., Kosarin S.P., Morozov S.P., Kuz'mina E.S. Unified radiological information service as a key element of organizational and methodical work of Research and Practical Center of Medical Radiology. *Radiologiya–Praktika (Radiology–Practice)*. 2018; 1 (67): 6–17 (in Russ.).]
7. Гусев А.В. Перспективы нейронных сетей и глубокого машинного обучения в создании решений для здравоохранения. *Врач и информационные технологии*. 2017; 3: 92–105. [Gusev A.V. Prospects for neural networks and deep machine learning in creating health solutions. *Vrach i Informatsionnye Tekhnologii (Information Technologies for the Physician)*. 2017; 3: 92–105 (in Russ.).]
8. Clark T.J., Flood T.F., Maximin S.T., Sachs P.B. Lung CT screening reporting and data system speed and accuracy are increased with the use of a semiautomated computer application. *J. Am. Coll. Radiol.* 2015; 12 (12 Pt A): 1301–6.
9. Барчук А.А., Арсеньев А.И., Беляев А.М., Гомболевский В.А., Нefeldова А.В., Канаев С.В. и др. Эффективность скрининга онкологических заболеваний. *Вопросы онкологии*. 2017; 4: 557–67. [Barchuk A.A., Arsen'ev A.I., Belyaev A.M., Gombolevskiy V.A., Nefedova A.V., Kanaev S.V. et al. The effectiveness of screening for cancer. *Voprosy Onkologii (Problems in Oncology)*. 2017; 4: 557–67 (in Russ.).]
10. Гомболевский В.А., Барчук А.А., Лайпан А.Ш., Ветшева Н.Н., Владимировский А.В., Морозов С.П. Организация и эффективность скрининга злокачественных образований легких методом низкодозной компьютерной томографии. *Радиология–практика*. 2018; 1 (67): 28–36. [Gombolevskiy V.A., Barchuk A.A., Laypan A.S., Vetsheva N.N., Vladimirovskiy A.V., Morozov S.P. Lung cancer screening with low-dose computed tomography: management and efficiency. *Radiologiya–Praktika (Radiology–Practice)*. 2018; 1 (67): 28–33 (in Russ.).]
11. Гомболевский В.А., Харламов К.А., Пятницкий И.А., Ким С.Ю., Морозов С.П. Шаблоны протоколов описаний исследований по специальности «рентгенология». Компьютерная томография. Методические рекомендации. 2016; 23: 13–4. URL: http://medradiology.moscow/d/1364488/d/metodicheskiye_rekomendatsii_no_23_2016_g_shablony_protokolov_opisaniy_kt_issledovaniy.pdf (дата обращения 20.10.2018). [Gombolevskiy V.A., Kharlamov K.A., Pyatnitskiy I.A., Kim S.Yu., Morozov S.P. Patterns of describing CT protocols examination. Guidelines. 2016; 23: 13–4 (in Russ.). Available at: http://medradiology.moscow/d/1364488/d/metodicheskiye_rekomendatsii_no_23_2016_g_shablony_protokolov_opisaniy_kt_issledovaniy.pdf (accessed 20 October 2018).]
12. Басманов С.Н., Басманова А.А. Обзор эволюции экспертных систем в медицине с точки зрения соответствия основным признакам. *Перспективы развития информационных технологий*. 2014; 21: 126–30. [Basmanov S.N., Basmanova A.A. A review of the evolution of expert systems in medicine in terms of compliance with the main features. *Perspektivy Razvitiya Informatsionnykh Tekhnologii*. 2014; 21: 126–30 (in Russ.).]
13. Bulu H., Sippo D.A., Lee J.M., Burnside E.S., Rubin D.L. Proposing new radlex terms by analyzing free-text mammography reports. *J. Digit. Imag.* 2018. DOI: 10.1007/s10278-018-0064-0
14. Goff D.J., Loehfelm T.W. Automated radiology report summarization using an open-source natural language processing pipeline. *J. Digit. Imag.* 2018; 31 (2): 185–92. DOI: 10.1007/s10278-017-0030-2
15. Huesch M.D., Cherian R., Labib S., Mahraj R. Evaluating report text variation and informativeness: natural language processing of CT chest imaging for pulmonary embolism. *J. Am. Coll. Radiol.* 2018; 15 (3 Pt B): 554–62. DOI: 10.1016/j.jacr.2017.12.017
16. Kahn C.E. Jr. An ontology-based approach to estimate the frequency of rare diseases in narrative-text radiology reports. *Stud. Health. Technol. Inform.* 2017; 245: 896–900.
17. Percha B., Zhang Y., Bozkurt S., Rubin D., Altman R.B., Langlotz C.P. Expanding a radiology lexicon using contextual patterns in radiology reports. *J. Am. Med. Assoc.* 2018; 10. DOI: 10.1093/jamia/ocx152
18. Pons E., Braun LM., Hunink M.G., Kors J.A. Natural language processing in radiology: a systematic review. *Radiology*. 2016; 279 (2): 329–43. DOI: 10.1148/radiol.16142770
19. Gálvez J.A., Pappas J.M., Ahumada L., Martin J.N., Simpao A.F., Rehman M.A. et al. The use of natural language processing on pediatric diagnostic radiology reports in the electronic health record to identify deep venous thrombosis in children. *J. Thromb. Thrombolysis*. 2017. DOI: 10.1007/s11239-017-1532-y
20. Tian Z., Sun S., Eguale T., Rochefort C.M. Automated extraction of VTE events from narrative radiology reports in electronic health records: a validation study. *Med. Care*. 2017; 55 (10): e73–e80. DOI: 10.1097/MLR.0000000000000346
21. Trivedi H., Mesterhazy J., Laguna B., Vu T., Sohn J.H. Automatic determination of the need for intravenous contrast in musculoskeletal MRI examinations using IBM Watson's natural language processing algorithm. *J. Digit. Imag.* 2017; 18. DOI: 10.1007/s10278-017-0021-3
22. Guimaraes C.V., Grzeszczuk R., Bisset G.S. 3rd, Donnelly L.F. Comparison between manual auditing and a natural language process with machine learning algorithm to evaluate faculty use of standardized reports in radiology. *J. Am. Coll. Radiol.* 2018; 15 (3 Pt B): 550–3. DOI: 10.1016/j.jacr.2017.10.042
23. Humphrey L.L., Deffebach M., Pappas M., Zakher B., Fu R., Slatore C.G. et al. Screening for lung cancer with lowdose computed tomography: a systematic review to update the U.S. Preventive services task force recommendation. *Ann. Int. Med.* 2013; 159 (4): 411–20.
24. Oudkerk M., Devaraj A., Vliegenthart R., Henzler T., Prosch H., Heussel C.P. European position statement on lung cancer screening. *Lancet Oncol.* 2017; 18 (12): 754–66.
25. Beyer S.E., McKee B.J., Regis S.M., McKee A.B., Flacke S., El Saadawi G. et al. Automatic Lung-RADS™ classification with a natural language processing system. *J. Thorac. Dis.* 2017; 9 (9): 3114–22. DOI: 10.21037/jtd.2017.08.13
26. Lacson R., Prevedello L.M., Andriole K.P., Gill R., Lenoci-Edwards J., Roy C. et al. Factors associated with radiologists' adherence to Fleischner Society guidelines for management of pulmonary nodules. *J. Am. Coll. Radiol.* 2012; 9 (7): 468–73. DOI: 10.1016/j.jacr.2012.03.009