

ИНСТИТУТ МОЗГА ЧЕЛОВЕКА ИМ. Н.П. БЕХТЕРЕВОЙ РОССИЙСКОЙ
АКАДЕМИИ НАУК

На правах рукописи

АНДРОПОВА ПОЛИНА ЛЕОНИДОВНА

ДИАГНОСТИКА ИШЕМИЧЕСКОГО ИНСУЛЬТА В БАССЕЙНЕ СРЕДНЕЙ
МОЗГОВОЙ АРТЕРИИ МЕТОДОМ КОМПЬЮТЕРНОЙ ТОМОГРАФИИ С
ПРИМЕНЕНИЕМ СИСТЕМ АВТОМАТИЧЕСКОГО АНАЛИЗА
ИЗОБРАЖЕНИЙ

Научная специальность 3.1.25. Лучевая диагностика

ДИССЕРТАЦИЯ

на соискание ученой степени

кандидата медицинских наук

Научный руководитель:

Член- корреспондент РАН, доктор медицинских наук, профессор

Трофимова Татьяна Николаевна

Санкт-Петербург – 2023

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ	4
ГЛАВА 1. ВОЗМОЖНОСТИ И ПЕРСПЕКТИВЫ ИНСТРУМЕНТАЛЬНОЙ ДИАГНОСТИКИ ИШЕМИЧЕСКОГО ИНСУЛЬТА (ОБЗОР ЛИТЕРАТУРЫ).....	19
1.1. Роль компьютерной томографии в диагностике ишемического инсульта в бассейне средней мозговой артерии.....	19
1.2. Шкала ASPECTS: проблемы межэкспертной вариативности в клинической практике	28
1.3. Возможности систем автоматического анализа компьютерно- томографических изображений в диагностике ишемического инсульта в бассейне средней мозговой артерии, а также в применении шкалы ASPECTS	37
ГЛАВА 2. ОБЩАЯ ХАРАКТЕРИСТИКА КЛИНИЧЕСКОГО МАТЕРИАЛА И МЕТОДЫ ИССЛЕДОВАНИЯ.....	50
2.1. Разработка базы данных компьютерно- томографических исследований пациентов с клинической картиной острого нарушения мозгового кровообращения в бассейне средней мозговой артерии	50
2.2. Проблема и вопросы межэкспертного согласия врачей рентгенологов с разным стажем и опытом работы в ургентной медицине при диагностике ишемического инсульта в бассейне средней мозговой артерии	59
2.3. Тестирование систем автоматического анализа	65
2.4. Комплементарная оценка компьютерно- томографических исследований системой автоматического анализа и врачами рентгенологами	73
ГЛАВА 3. РЕЗУЛЬТАТЫ КТ- ДИАГНОСТИКИ ИШЕМИЧЕСКОГО ИНСУЛЬТА В БАССЕЙНЕ СРЕДНЕЙ МОЗГОВОЙ АРТЕРИИ ВРАЧАМИ РЕНТГЕНОЛОГАМИ В ЗАВИСИМОСТИ ОТ СТАЖА И ОПЫТА РАБОТЫ В УРГЕНТНОЙ МЕДИЦИНЕ	77

ГЛАВА 4. РЕЗУЛЬТАТЫ ТЕСТИРОВАНИЯ СИСТЕМ АВТОМАТИЧЕСКОГО АНАЛИЗА И ИХ ВЛИЯНИЕ НА ПРОЦЕСС ПРИНЯТИЯ ВРАЧЕБНОГО РЕШЕНИЯ	90
ГЛАВА 5. ВЛИЯНИЕ ВНЕДРЕНИЯ СИСТЕМ АВТОМАТИЧЕСКОГО АНАЛИЗА НА ПРОЦЕСС ПРИНЯТИЯ ВРАЧЕБНОГО РЕШЕНИЯ	98
ЗАКЛЮЧЕНИЕ	113
ВЫВОДЫ	125
ПРАКТИЧЕСКИЕ РЕКОМЕНДАЦИИ	127
ПЕРСПЕКТИВЫ ДАЛЬНЕЙШЕЙ РАЗРАБОТКИ ТЕМЫ	128
СПИСОК СОКРАЩЕНИЙ	130
СЛОВАРЬ ТЕРМИНОВ	132
СПИСОК ЛИТЕРАТУРЫ	134

ВВЕДЕНИЕ

Актуальность темы исследования

По экспертным оценкам Всемирной организации здравоохранения (ВОЗ), ишемический инсульт является важнейшей медико-социальной проблемой, обусловленной его высокой долей в структуре заболеваемости и смертности населения, значительными показателями временных трудовых потерь и первичной инвалидности [62]. По данным Федеральной службы государственной статистики Российской Федерации (РФ), цереброваскулярные заболевания занимают второе место в структуре смертности от сердечно-сосудистых заболеваний (39%), доля ОНМК составляет 21,4%. В острый период ишемического инсульта летальность достигает 35%, и к первому году с момента развития заболевания умирают 50% больных. В период с 2008 по 2016 г. смертность от инсульта в Российской Федерации снизилась на 45% и составила 123 случая на 100 тыс. населения, в 2019 г. этот показатель достиг предельно низких результатов — 88,2 случая на 100 тыс. населения. Однако на фоне эпидемии новой коронавирусной инфекции, этот показатель за 2020 г. возрос до 92,4 случаев на 100 тыс. населения и продолжает расти [12, 13].

На сегодняшний день в диагностике острого нарушения мозгового кровообращения (ОНМК) нейровизуализация занимает ведущую позицию. Согласно принятым в РФ клиническим рекомендациям, всем пациентам при подозрении на ишемический инсульт рекомендуется проведение бесконтрастной КТ или МРТ головного мозга в экстренном порядке, с получением результатов исследования в течение 40 минут от момента поступления в стационар для дифференциальной диагностики формы ОНМК и определением тактики лечения [12, 23]. С учетом ряда причин (время проведения исследования, отсутствия абсолютных противопоказаний, доступность), компьютерная томография (КТ) является ведущим методом нейровизуализации в диагностике ОНМК, который

призван первоначально ответить на два основных вопроса: наличие ишемического поражения головного мозга и/ или внутримозгового кровоизлияния [12].

Распознавание КТ- признаков ишемии на ранних стадиях является сложной диагностической задачей. С целью формирования единого подхода к диагностике ишемического инсульта, в 2000 г. была разработана методика ASPECTS (Alberta stroke program early CT score) – полуколичественная шкала, созданная для оценки распространенности ранних ишемических изменений в бассейне средней мозговой артерии по 10 – бальной системе. Оценка ≤ 7 баллов сообщает о более выраженном объеме поражения мозговой ткани в бассейне СМА и коррелирует с худшим функциональным исходом, а также более высоким риском геморрагической трансформации инсульта [3, 28].

Шкала ASPECTS получила положительную оценку многих исследователей как надежный диагностический метод. Однако использование данной 10-балльной системы имеет и ряд ограничений, куда следует отнести отсутствие стандартизации, приводящей к высокой вариативности межэкспертной оценки ишемических изменений с помощью ASPECTS, что может отражаться на дальнейшем ходе лечения пациента [3, 75].

Для частичного решения проблемы субъективности применения шкалы ASPECTS предлагается внедрение систем автоматического анализа в качестве метода поддержки принятия врачебных решений. На сегодняшний день данные системы разрабатываются для анализа бесконтрастных КТ, КТ- ангиографии и КТ- перфузии. Их применение направлено на определение балла по шкале ASPECTS, количественную оценку ядра инсульта, пенумбры, статуса коллатерального кровотока и локализации артериальной окклюзии в автоматическом режиме [3, 5, 113].

Использование алгоритмов искусственного интеллекта (ИИ) предполагает более эффективное обнаружение ишемических изменений и снижение межэкспертной вариативности при оценке КТ- изображений пациентов, нуждающихся в экстренной медицинской помощи. Развитие способности

сочетать коллективные рассуждения и прогнозы систем автоматического анализа может оказать глубокое влияние на организацию и управление медицинской помощью [49, 113].

Таким образом, актуальным вопросом является валидация разработанных алгоритмов искусственного интеллекта и их внедрение в клиническую практику, с целью поддержки принятия врачебных решений и стандартизации интерпретации данных КТ- исследований, что может улучшить тактику лечения пациентов и функциональный исход ишемического инсульта.

Степень разработанности темы

Чувствительный ко времени характер оказания помощи при инсульте («Время – Мозг») обуславливает потребность в точных и быстрых инструментах помощи при диагностике инсульта [12, 63]. За последние годы одной из тенденций развития клинической медицинской практики и темой исследовательских работ международного научного сообщества является разработка и внедрение алгоритмов автоматической оценки изображений нейровизуализации на основе искусственного интеллекта [31]. Были представлены различные автоматизированные системы определения объема распределения ишемических изменений в бассейне СМА по шкале ASPECTS, размера пенумбры и ядра инсульта по данным КТ- перфузии, а также выявления тромбоза сосудов головного мозга, согласно данным КТ – ангиографии [51, 53]. Различные зарубежные IT- компании анонсируют внедрение коммерчески доступных автоматизированных и полуавтоматических программных обеспечений для диагностики острого инсульта в стандартный рабочий процесс (Aidoc®, Apollo Medical Imaging Technology®, Brainomix®, inferVISION®, RAPID®, JLK Inspection®, Max-Q AI®, Nico.lab®, Olea Medical®, Qure.ai®, Viz.ai® и Zebra Medical Vision®).

В настоящее время, согласно утвержденной Национальной стратегии развития искусственного интеллекта на период до 2030 г., утвержденной указом

№ 490 от 10 октября 2019 г., в РФ активно идут разработки систем автоматического анализа, в том числе для применения в сфере здравоохранения и в нейрорадиологии, в частности [9, 21]. В качестве инструмента регулирования данной деятельности, был создан подкомитет "Искусственный интеллект в здравоохранении" на базе Центра диагностики и телемедицины Департамента здравоохранения города Москвы в 2019 г. В рамках деятельности данного подкомитета, в Российской Федерации осуществляется разработка национальных стандартов [17, 18], а также клинические рекомендации по испытанию программного обеспечения на основе интеллектуальных технологий [15]. Одним из направлений по внедрению в здравоохранение РФ систем поддержки принятия врачебных решений является диагностика ишемического инсульта. Главной задачей, поставленной перед разработчиками алгоритмов искусственного интеллекта, является достижение их системами пороговых значений эффективности, не только определенных в клинических рекомендациях [15], но и результативности оценки сопоставимой с показателями молодых специалистов, к примеру, врачами рентгенологами с небольшим стажем (до 3 лет).

Использование систем автоматического анализа компьютерно-томографических изображений потенциально предполагает сокращение числа случаев гипо- и гипердиагностики, а также уменьшение межэкспертной вариативности при распознавании ишемического инсульта с нивелированием человеческого фактора [3, 89, 113]. Несмотря на высокую перспективность их практического применения, в литературе отсутствуют публикации, посвященные независимой оценке таких систем в структуре КТ- диагностики ишемического инсульта.

С учетом вышеизложенной проблематики существует потребность изучения диагностических возможностей систем искусственного интеллекта посредством проведения аналитической и клинической валидации, с исследованием особенностей взаимодействия врача-рентгенолога и систем автоматического анализа изображений.

Цель исследования

Совершенствование диагностики ишемического инсульта в бассейне средней мозговой артерии методом рентгеновской компьютерной томографии путем использования систем автоматического анализа изображений.

Задачи исследования

Для достижения поставленной цели были определены следующие задачи:

1. Сформировать базу данных, включающую анонимизированные компьютерно - томографические исследования пациентов с верифицированным острым нарушением мозгового кровообращения и без патологических изменений головного мозга.
2. Оценить показатели диагностической эффективности врачей рентгенологов с разным стажем и опытом работы в ургентной медицине при диагностике ишемического инсульта в бассейне средней мозговой артерии, а также межэкспертного согласия специалистов неотложной нейрорадиологии при оценке ишемических изменений по ASPECTS.
3. Оценить возможности применения систем автоматического анализа компьютерно- томографических изображений как метода выявления ишемического инсульта в бассейне средней мозговой артерии.
4. Разработать оптимальный алгоритм применения систем автоматического анализа компьютерно- томографических изображений в практике врача рентгенолога как метода выявления ишемического инсульта в бассейне средней мозговой артерии.
5. Обосновать рекомендации по выбору модели взаимодействия врача-рентгенолога и системы автоматического анализа изображений.

Научная новизна

Доказана зависимость показателей диагностической эффективности при выявлении ишемического инсульта в бассейне средней мозговой артерии от стажа врача - рентгенолога и наличия у него опыта работы в ургентной медицине.

Выявлены низкие показатели воспроизводимости шкалы ASPECTS среди врачей региональных сосудистых центров вне зависимости от стажа работы.

Доказана значимость выбора модели применения систем автоматического анализа КТ- изображений при комплементарной оценке ишемических изменений с врачами рентгенологами.

Доказано, что несмотря на низкие показатели точности (менее 0,8), согласно клиническим рекомендациям [15], при комплементарной оценке система автоматического анализа способствует увеличению диагностической эффективности врачей рентгенологов.

Выявлена положительная корреляция снижения вариативности оценки ишемических изменений в бассейне средней мозговой артерии по ASPECTS между молодыми специалистами со стажем до 3 лет при внедрении системы автоматического анализа.

Теоретическая и практическая значимость

В ходе диссертационной работы подтверждена прямая зависимость диагностической эффективности при выявлении ишемического инсульта в бассейне средней мозговой артерии с помощью компьютерной томографии от опыта и стажа работы в ургентной медицине врачей рентгенологов. Установлен низкий уровень межэкспертного согласия при оценке ишемических изменений по ASPECTS врачами с различным стажем, специализирующимися на диагностике ОНМК. Обоснована целесообразность рассмотрения системы автоматического анализа в качестве второго мнения для врачей-рентгенологов со стажем до трех лет. На основании исследования разработаны практические рекомендации по выбору модели взаимодействия врача-рентгенолога и системы

автоматического анализа для комплементарной оценки ишемических изменений в бассейне средней мозговой артерии с помощью компьютерной томографии.

Методология и методы исследования

Диссертационное исследование выполнялось в несколько этапов. На первом этапе был проведен детальный анализ литературы, посвященной данной проблеме. Обзор литературы основан на 118 источниках из них 23 - отечественных, 95 - зарубежных.

На втором этапе была сформирована и зарегистрирована база данных на основе результатов исследований методом рентгеновской компьютерной томографии (КТ) 150 пациентов с клинической картиной ОНМК в бассейне средней мозговой артерии (свидетельство о регистрации базы данных RU 2022620850). В базу данных вошли 100 пациентов с подтвержденным ишемическим инсультом в бассейне средней мозговой артерии с помощью КТ-ангиографии и КТ-перфузии, а также 50 пациентов с исключенным ОНМК, на основе динамического КТ-наблюдения и КТ-ангиографии. Были сформированы две выборки нативных КТ-изображений для дальнейшего тестирования врачей-рентгенологов и систем автоматического анализа компьютерно-томографических изображений. Также были составлены 3 варианта выборок для дальнейших этапов исследований: 1А и 2А для третьего этапа, 1В, 1С, 2В и 2С для пятого этапа исследования (Рисунок 1).

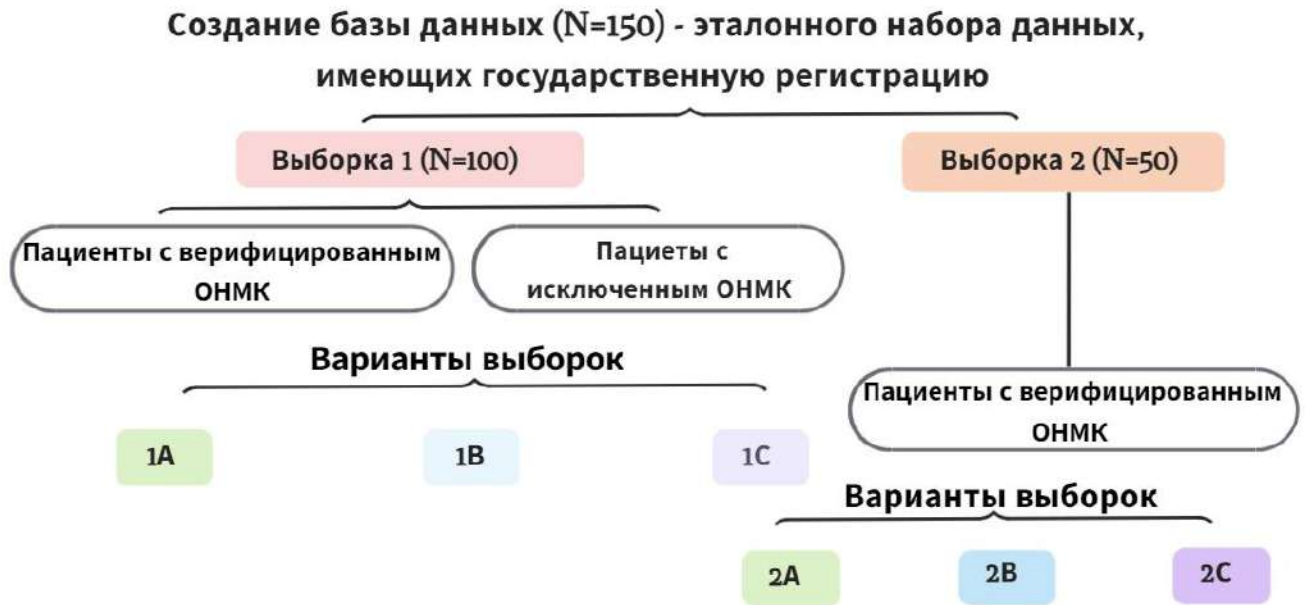


Рисунок 1 – Схематическое изображение 2 этапа исследования с иллюстрацией формирования выборок и из вариаций (1А, 2А для 3 этапа исследования, 1В, 1С, 2В и 2С для пятого этапа исследования)

На третьем этапе диссертационного исследования были протестированы врачи- рентгенологи с различным стажем и наличием опыта в диагностике инсульта на предмет выявления показателей эффективности в определении ОНМК в бассейне средней мозговой артерии с помощью компьютерной томографии, а также оценки межэкспертной согласованности в отношении КТ-признаков ишемического инсульта (гиперденсная средняя мозговая артерия, нарушение дифференцировки белого и серого вещества, сглаженность корковых борозд и снижение КТ- плотности вещества головного мозга). Также был выявлен уровень межэкспертной согласованности при определении распространенности ишемических изменений по ASPECTS врачами, специализирующимися в неотложной нейрорадиологии с разным стажем работы (Рисунок 2).

Вариант выборки 1А



Вариант выборки 2А



Рисунок 2 – Схематическое изображение 3 этапа исследования, проведенного в период 23.05 – 03.06 2022, на выборках 1А и 2А, с тестированием врачей с различным стажем из региональных сосудистых центров (РСЦ) и из стационаров общего профиля (СОП)

На четвертом этапе было проведено тестирование систем автоматического анализа методом аналитической валидации, с определением показателей их диагностической точности. После выбора программы для дальнейших исследований, она была протестирована на выборке КТ-изображений пациентов с подтвержденным инсультом в бассейне средней мозговой артерии для определения объема ишемических изменений по ASPECTS (Рисунок 3).

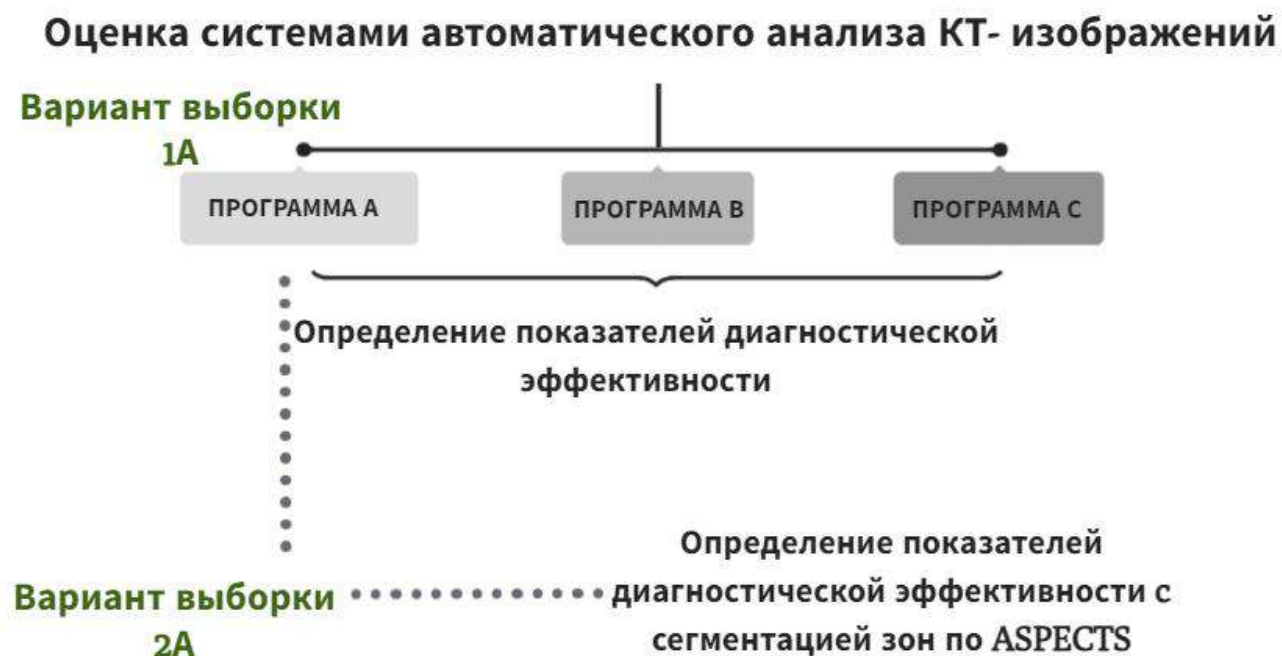


Рисунок 3 – Схематическое изображение 4 этапа исследования с тестированием анонимизированных систем автоматического анализа (программы А, В, С)

На пятом этапе были изучены возможные варианты внедрения систем автоматического анализа компьютерно- томографических изображений в качестве инструмента поддержки врачебных решений рентгенологов со стажем менее 3 лет посредством проведения тестирования с моделированием двух различных вариантов комплементарной оценки: моделей первого (параллельного) и второго чтения (Рисунок 4).

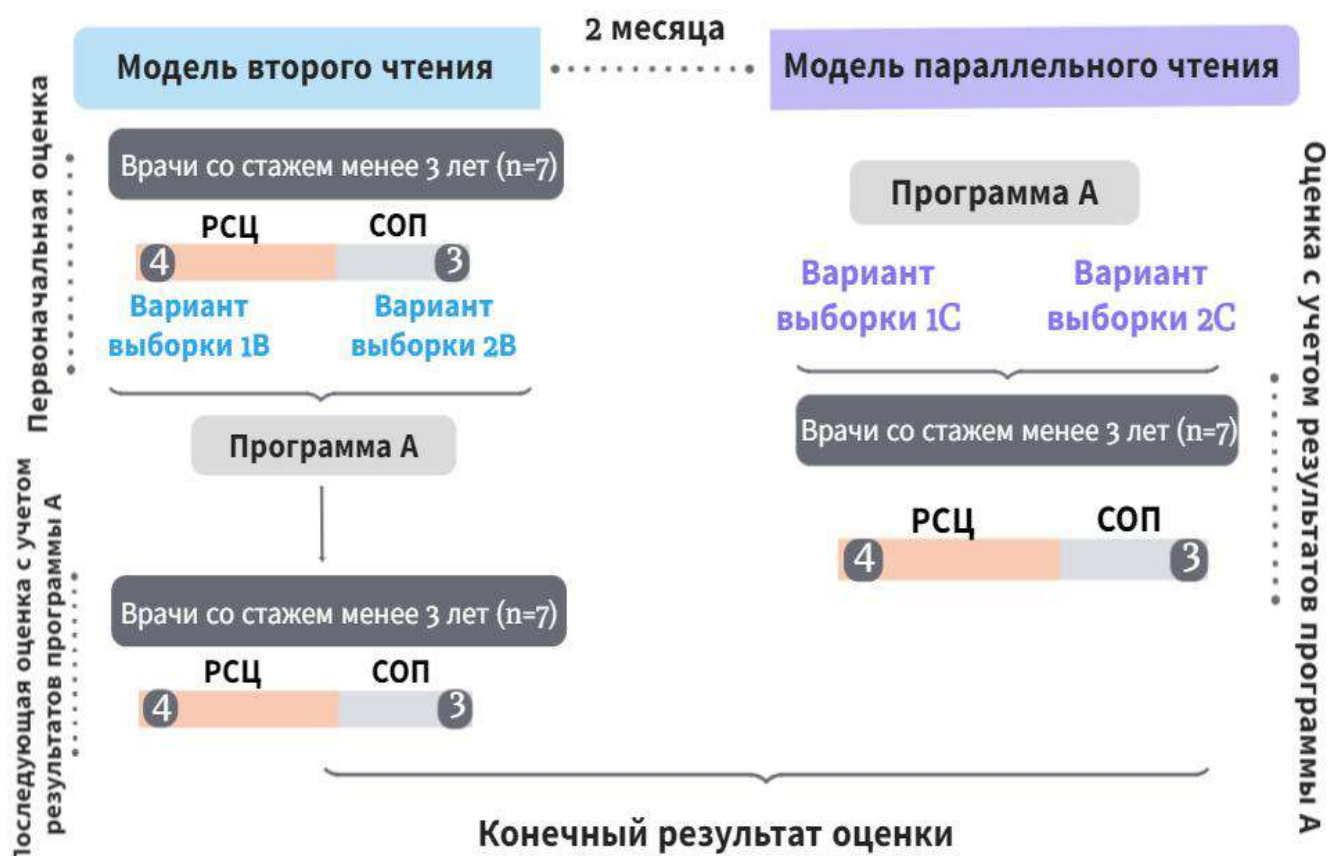


Рисунок 4 – Схематическое изображение 5 этапа исследования, проведенного в период 03.05-05.05.2023, на выборках 1В и 2В и в период 01.07-04.07.2023, на выборках 1С и 2С

Статистическую обработку проводили в программе «Statistical Package for the Social Sciences» (SPSS), версия 19.0 (SPSS Inc., Chicago, IL, USA). Проводилась оценка показателей эффективности выявления ишемического инсульта в бассейне средней мозговой артерии (истинно положительные, ложноположительные, ложноотрицательные, истинно отрицательные результаты, чувствительность, специфичность, точность), а также уровня межэкспертного согласия (каппа статистика: каппа Коэна и каппа Флейсса). Проведен ROC-анализ и рассчитана площадь под ROC-кривой.

Положения, выносимые на защиту

1. Шкала ASPECTS имеет низкие показатели воспроизводимости при оценке ишемических изменений врачами вне зависимости от стажа и наличия опыта в ургентной медицине.
2. Автоматический анализ компьютерных томограмм с использованием специального программного обеспечения позволяет повысить показатели диагностической эффективности врачей рентгенологов со стажем менее 3 лет с помощью применения модели параллельного чтения.
3. Автоматический анализ компьютерных томограмм с использованием специального программного обеспечения позволяет повысить согласованность врачей рентгенологов со стажем менее 3 лет при оценке ишемических изменений в бассейне средней мозговой артерии по ASPECTS с помощью применения модели параллельного чтения.

Степень достоверности и апробации результатов

Степень достоверности результатов проведенного исследования определяется достаточным количеством клинических наблюдений, репрезентативным объемом выборки, применением современной методологии тестирования врачей-рентгенологов с различным стажем и опытом работы (выявление признаков инсульта в бассейне средней мозговой артерии и оценки ишемических изменений по ASPECTS), проведением тестирования систем автоматического анализа, а также совместного тестирования сопоставимых по стажу работы врачей-рентгенологов системы автоматического анализа и обработкой полученных данных адекватными методами математической статистики.

Материалы диссертации представлены на конференциях

Основные результаты, полученные в рамках работы над диссертацией, были представлены на VIII-ой Конгрессе Национальной Ассоциации Фтизиатров (Санкт-Петербург, секция «СИМПОЗИУМ Искусственный интеллект в лучевой диагностике», 25 – 27 ноября 2019 г), XII Международном конгрессе «Невский радиологический Форум – 2021» (Санкт- Петербург, 7–10 апреля 2021 г), Всероссийском национальном конгрессе лучевых диагностов и терапевтов «Радиология – 2021» (Москва, 25–27 мая 2021 г), European Congress of Radiology (ECR) / Конгрессе Европейского общества радиологов (Вена, 2-6 марта 2022 г), European Congress of Radiology (ECR) / Конгрессе Европейского общества радиологов (Вена, 13-17 июля 2022 г), Joint Workshop” Machine Learning Methods and Statistical Models in Medicine” Saint Petersburg State University (SPSU) – Huazhong University of Science and Technology (HUST) (Санкт- Петербург, 29 сентября 2022 г), XIV Международном конгрессе «Невский радиологический Форум – 2021» (Санкт- Петербург, 7–8 апреля 2023 г .), XVII Всероссийском национальном конгрессе лучевых диагностов и терапевтов «Радиология – 2023» (Москва, 30 мая - 1 июня 2023 г), конференции «Вычислительная биология и искусственный интеллект для персонализированной медицины-2023» (он-лайн, 9-11 августа 2023 г), VI Съезде Национального Общества нейрорадиологов (Сочи, 29-30 сентября 2023 г).

Практическая реализация работы

Результаты исследования и разработки внедрены в практическую деятельность рентгеновского отделения СПб ГБУЗ «Елизаветинская больница», рентгеновского отделения Краевой больницы №1 им. Очаповского, рентгеновского отделения Самарской областной клинической больницы им. В. Д. Середавина, в многопрофильной медицинской клинике для детей и взрослых "Скандинавия" в Санкт-Петербурге. Полученные теоретические и практические результаты диссертационной работы используются в учебном процессе ФГБУН

ИМЧ им. Н.П. Бехтеревой РАН, ФГБУ «СПб НИИФ» Минздрава России и ФГБОУ ВО "КубГМУ" Минздрава России.

Публикации

По теме диссертационного исследования опубликовано 6 печатных работ, из них 3 публикации в изданиях, рекомендованных Высшей аттестационной комиссией Министерства образования и науки Российской Федерации для опубликования результатов диссертационных работ, 2 публикации в изданиях, относящихся к Scopus, а также зарегистрирована база данных (Свидетельство о регистрации № 2022620850). Научные публикации достаточно полно отражают содержание диссертации и автореферата.

Личный вклад автора в получении результатов

Тема и план диссертации, ее основные идеи и содержание разработаны совместно с научным руководителем на основе детального изучения литературы. Автор самостоятельно сформулировала и обосновала актуальность темы диссертации, цель, задачи и этапы научного исследования. Диссертант лично изучила литературу и совместно с научным руководителем разработала методологию исследования. Диссертант является со-разработчиком базы данных лучевых изображений, на основе которых проводилась аналитическая валидация программ автоматического анализа, а также тестирование врачей - рентгенологов. Диссертант лично провел все исследования, вошедшие в базу данных, а также тестирование всех выбранных программ для анализа изображений, с последующей интерпретацией полученных результатов, собрал и проанализировал комплексные данные компьютерной томографии. Личный вклад автора в изучение литературы, сбор, обобщение, анализ, статистическую обработку полученных данных и написание диссертации – 100%. Автором лично написан текст диссертации.

Соответствие диссертации паспорту специальности

Диссертация соответствует паспорту специальности 3.1.25. «Лучевая диагностика»

п. 1. Диагностика и мониторинг физиологических и патологических состояний, заболеваний, травм и пороков развития (в том числе внутриутробно) путем оценки качественных и количественных параметров, получаемых с помощью методов лучевой диагностики.

п. 11. Использование цифровых технологий, искусственного интеллекта и нейросетей для диагностики и мониторинга физиологических и патологических состояний, заболеваний, травм и пороков развития (в том числе внутриутробно) с помощью методов лучевой диагностики.

Объем и структура диссертации

Диссертация представлена на 148 страницах текста компьютерного набора формата New Times Roman, шрифт 14. Диссертационное исследование состоит из введения, обзора литературы, описания материалов и методов исследования, глав с результатами собственных исследований, заключения, выводов, практических рекомендаций, перспектив дальнейшей разработки темы, списка литературы. Работа иллюстрирована 28 таблицами и 18 рисунками.

ГЛАВА 1. ВОЗМОЖНОСТИ И ПЕРСПЕКТИВЫ ИНСТРУМЕНТАЛЬНОЙ ДИАГНОСТИКИ ИШЕМИЧЕСКОГО ИНСУЛЬТА (ОБЗОР ЛИТЕРАТУРЫ)

1.1. Роль компьютерной томографии в диагностике ишемического инсульта в бассейне средней мозговой артерии

По экспертным оценкам Всемирной организации здравоохранения (ВОЗ), инсульт занимает второе место в мире среди причин смертности. Ежегодно прогнозируется развитие 9,6 миллионов инсультов с ростом заболеваемости по мере старения населения, при этом 85% случаев приходится на долю ишемического инсульта [62].

В Российской Федерации в период с 2008 по 2016 г. смертность от инсульта снизилась на 45% и составила 123 случая на 100 тыс. населения, в 2019 г. этот показатель достиг предельно низких результатов — 88,2 случая на 100 тыс. населения. Однако на фоне эпидемии новой коронавирусной инфекции, этот показатель за 2020 г возрос до 92,4 случаев на 100 тыс. населения и продолжает расти [12].

Показатели смертности у больных с инсультом в значительной степени зависят от условий лечения в остром периоде. Показатель ранней 30-дневной смертности на фоне ОНМК достигает 35%. При этом, в стационарах умирает около 24% пациентов, в домашних условиях – 43%, к первому году с момента развития заболевания погибают 50% больных [1, 23].

Наравне с высокой смертностью к социально значимым последствиям ОНМК относится инвалидизация выживших больных после первичного инсульта, а также повышенный риск развития повторных инсультов. По данным Национального регистра приблизительно 60% больных, перенесших инсульт остаются инвалидами, способными себя обслуживать, 19-35% нуждаются в посторонней помощи для ухода за собой и становятся зависимыми от

окружающих, и только 15-20% сохраняют трудовую активность. Согласно данным эпидемиологического исследования инсульта методом территориально-популяционного регистра, отношение ишемических инсультов к геморрагическим составило 5:1. Средний возраст развития инсульта - 66,7 лет (63,7 года у мужчин и 69,4 года у женщин). Абсолютное число инсультов у пациентов в возрасте до 67 лет выше у мужчин, а в более старшем возрасте выше у женщин [23].

Необходимо отметить медико-экономическую значимость проблемы инсульта. По сведениям ВОЗ, затраты на лечение и реабилитацию одного пациента с ОНМК составляют 55 000-73000 долларов/год, что свидетельствует об огромном экономическом ущербе от этого заболевания [63].

Анализ вышеизложенных данных показывает, что лечение пациентов с цереброваскулярной патологией – одна из самых актуальных и сложных проблем медицины и социальной помощи, обусловленная стремительно растущей заболеваемостью, а также крайне тяжелыми последствиями, приводящими к высокой инвалидизации.

Самой часто поражаемой внутримозговой артерией при ишемическом инсульте является средняя мозговая артерия (СМА). Она ответвляется непосредственно от внутренней сонной артерии и состоит из четырех основных сегментов: М1, М2, М3 и М4. Эти сосуды обеспечивают кровоснабжение большей части лобных, височных и теменных долей головного мозга, а также (лентикулостриарные артерии) более глубоких структур полушарий, в том числе базальные ядра [14, 85], что обуславливает худший функциональный исход и более высокий риск летальности при неадекватном лечении [4, 85]. Потому усовершенствование диагностики ишемического инсульта в бассейне средней мозговой артерии имеет такую высокую значимость.

В диагностике инсульта ведущее место принадлежит таким методам визуализации, как КТ и МРТ [22, 60, 86, 117]. КТ является высокоинформативным методом, с помощью которого возможно выявление ОНМК и его характера (ишемический, геморрагический, смешанный тип) [55]. Метод позволяет

определять локализацию, размеры (объем) и характер очага поражения с высокой точностью, которая при исследовании некоторых видов нарушений мозгового кровообращения достигает 90-100% [36]. Повторное КТ-сканирование дает возможность проследить эволюцию патологических изменений во времени.

Согласно клиническим рекомендациям [12] всем пациентам с признаками ОНМК в экстренном порядке рекомендуется проведение бесконтрастной КТ или МРТ головного мозга с получением результатов исследования (заключения) в течение 40 минут от момента поступления пациента в стационар для дифференциальной диагностики ОНМК с целью определения тактики лечения [4, 10, 56, 103, 104, 110, 118].

Уменьшение интервала времени с момента поступления в стационар до начала нейровизуализационного исследования головного мозга способствует сокращению времени до начала терапии и, соответственно, лучшему клиническому исходу заболевания.

В соответствии с постановлениями Правительства РФ от 27 декабря 2007 г. № 1012 и № 186 от 2 марта 2009 г. в РФ начата реализация мероприятий, направленных на совершенствование медицинской помощи больным сосудистыми заболеваниями. В составе региональных сосудистых центров и первичных сосудистых отделений созданы подразделения для лечения пациентов с ОНМК, с круглосуточно функционирующими подразделениями лучевой диагностики (кабинета КТ и/или МРТ), а также отделением (кабинетом) рентгенохирургических методов диагностики, лечения и операционной рентгенэндоваскулярных методов диагностики и лечения. Одним из непредвиденных последствий данной инициативы стала госпитализация практически всех пациентов с нарушением сознания различной этиологии с диагнозом ОНМК в сосудистый центр сотрудниками бригад скорой медицинской помощи. Следствием такого подхода скорой медицинской помощи стало массовое поступление пациентов в первичные сосудистые отделения и региональные сосудистые центры с подозрением на ОНМК, обуславливающее повышение

нагрузки на специалистов, участвующих в экстренной диагностике инсульта на стационарном этапе (неврологи и рентгенологи), а также повышение использования серьезных материальных ресурсов [95]. Также, пациентам с внезапно развившимся ОНМК, находящимся в стационарах, оказывающих медицинскую помощь по основному заболеванию, выполняют КТ для исключения другой этиологии клинической картины, для обоснования перевода в сосудистый центр.

В условиях отсутствия разделения лучевых диагностов по специализации, по примеру зарубежной медицинской практики, каждый специалист компьютерной томографии в РФ должен демонстрировать высокую диагностическую эффективность в обнаружении ишемического инсульта, что невозможно без определенных компетенций, опыта и практических навыков работы с указанным контингентом пациентов.

Ранняя КТ диагностика ОНМК дает возможность выявить и оценить ишемические изменения, что, в свою очередь, определяет тактику лечения и прогноз осложнений [4, 80].

Существует ряд КТ-признаков, которые указывают на наличие ишемического инсульта: признак острой окклюзии артерии (гиперденсивность средней мозговой артерии и ее ветвей), признаки ранних ишемических изменений (нарушение дифференцировки серого/белого вещества, сглаженность корковых борозд на стороне поражения), а также снижение рентгеновской плотности вещества головного мозга [4, 98].

Гиперденсивность средней мозговой артерии, как высокоспецифичный признак (90-100%) острого ишемического инсульта в одноименном бассейне, был впервые описан Gacs G. и соавт. в 1983 г. [47]. Гиперденсивность СМА проявляется повышением КТ-плотности проксимальных отделов СМА в достаточно широком диапазоне от +40 HU до +80 HU, и ассоциирована с тромбозом M1 (и зачастую M2) сегмента. Такой же механизм определяется в

дистальных отделах СМА с формированием картины точечного повышения плотности СМА (“признак точки”) [4, 68, 114].

Гиперденсивность СМА нередко бывает единственным диагностическим КТ-признаком на ранней стадии ишемического инсульта [4].

Однако не стоит забывать, что несмотря на свою высокую специфичность, гиперденсивность СМА может встречаться при повышенном гематокрите (в отличие от данного признака, повышение КТ-плотности будет наблюдаться у всех внутричерепных артерий и вен), при вирусных поражениях [4, 76], также известны упоминание гиперденсивности СМА при диссекции артерии [4, 66]. Снижение плотности, окружающей артерию паренхимы мозга (инфекционное/неопластическое поражение) также могут создавать картину псевдогиперденсивного сосуда [4, 74].

Коо С. К. и соавт. [76] установили объективные критерии гиперденсивности СМА путем измерения абсолютной плотности пораженных и неизмененных сосудов: абсолютная плотность должна составлять не менее +43 НУ при соотношении с плотностью других артерий $<1,2$ [4].

Отек является распространенным ответом на различные формы повреждения головного мозга. При ишемическом инсульте выделяют три вида отека: цитотоксический, ионный и вазогенный. Каждый из этих видов обладает своим доминирующим паттерном визуализации, идентификация которого в сочетании с дополнительными лучевыми данными и клиническим анамнезом, часто дает ключ к верной трактовке изменений и влияет на дальнейшее лечение [4, 39].

Цитотоксический отек возникает в результате нарушения зависимых от аденозинтрифосфата (АТФ) трансмембранных натрий-калиевых и кальциевых насосов, и обычно вызывается церебральной ишемией или эксайтотоксическим (вторичным по отношению к чрезмерной нейромедиаторной стимуляции) повреждением головного мозга. Это приводит к внутриклеточному накоплению жидкости в нейронах и нейроглии. В первую очередь поражается серое вещество

из-за его высокой метаболической активности и большей плотности астроцитов [4, 98]. На данном этапе развития ишемического процесса, КТ не является чувствительным методом его обнаружения, а визуализационный паттерн возможно выявить только с помощью МРТ [4].

Далее, после исчерпания всех компенсаторных механизмов, по осмотическому градиенту вода начинает поступать в истощенное межклеточное пространство, но без повреждения гематоэнцефалического барьера, моделируя ионный отек. При данном процессе в конечном итоге вовлекаются как серое, так и белое вещество с соответствующей потерей их дифференцировки на КТ [4, 39].

Нарушение дифференцировки коры и белого вещества, в первую очередь, включает в себя «признак инсулярной ленты» (утрата ребристой структуры извилин в области островка) и исчезновение границ (контрастности) базальных ганглиев, а в дальнейшем, изменение дифференцировки коры больших полушарий [74]. Раннее вовлечение базально-ядерной области в патологический процесс при тромбозе проксимальных отделов СМА обуславливается тем, что данная анатомическая область снабжается перфорирующими (латеральными лентикулостриарными) артериями, ответвляющимися непосредственно от М1 сегмента СМА [4].

Тяжелые или повторяющиеся повреждения перегружают трансмембранные ионные насосы, вызывая гибель клеток с нарушением гематоэнцефалического барьера, и, в результате, возникает вазогенный отек. Поздние осложнения включают апоптоз нейронов, атрофию и глиоз [4].

Вазогенный отек моделируется разрушением плотных эндотелиальных соединений, образующих гематоэнцефалический барьер, вторично по отношению либо к физическому разрушению, либо к высвобождению вазоактивных медиаторов. В результате внутрисосудистые белки и жидкость поступают во внеклеточное пространство [4, 39].

Увеличение содержания жидкости в межклеточном пространстве на 1% приводит к снижению КТ-плотности головного мозга на 2,5 единицы

Хаунсфилда. С учетом того, что вазогенный отек приводит к нарушению гематоэнцефалического барьера, повышенная проницаемость сосудов может привести к геморрагической трансформации инсульта и экстравазации контрастного препарата после внутривенной тромболитической терапии и/или эндоваскулярного вмешательства. И чем больше зона изменений, тем выше эта вероятность. Именно с целью оценки риска подобных осложнений была введена шкала ASPECTS (полуколичественная шкала, созданная для оценки ранних ишемических изменений в бассейне средней мозговой артерии по 10- балльной системе), с помощью которой на основе всех вышеописанных признаков можно спрогнозировать функциональный исход инсульта, а также результаты тромболитической терапии и тромбэкстракции [4, 57].

Сглаженность корковых борозд описывается как следствие масс- эффекта, которая может свидетельствовать о необратимом повреждении паренхимы головного мозга [4, 36]. Однако Haussen D. C. и соавт. в своем исследовании подтвердили, что несмотря на общепринятую трактовку значения данного признака, сглаженность борозд без ассоциированного снижения плотности вещества головного мозга, с сохранением дифференцировки коры не может интерпретироваться как признак ишемических изменений, а также учитываться при оценке по ASPECTS, поскольку является потенциально обратимым явлением, за счет повышенного лептоменингеального коллатерального кровотока [4, 105]. С учетом отсутствия единого подхода в определении данного признака, по данным зарубежной литературы, он может вызвать вариативность результатов оценки КТ-признаков ишемического инсульта [4].

В статье Wardlaw J. M. и соавт. [116] был проведен систематический обзор публикаций, выпущенных между 1990 и 2003 гг., содержащих исследования оценки согласованности между экспертами в выявлении признаков ранних ишемических изменений (в т.ч. последующая оценка по ASPECTS), и определения корреляции между результатами диагностического анализа, клиническим исходом, а также лечением (тромболизис). По данным 15

исследований согласованность между экспертами (в среднем 30 КТ исследований и шесть экспертов) в отношении наличия РИИ варьировалась между 0,14 и 0,78 для любого РИИ (каппа статистика). Средняя чувствительность и специфичность выявления ранних ишемических изменений составила 66% (диапазон 20–87%) и 87% (диапазон 56–100%) соответственно. При исследованиях, в которых были задействованы в качестве нейрорадиологии с большим стажем и экспертной квалификацией, точность выявления РИИ возрастала. В статьях, где диагностическую эффективность пытались повысить за счет однородной и общей теоретической подготовки экспертов, результаты после проводимого эксперимента не улучшились. Авторы пришли к выводу, что необходимы дальнейшие исследования для определения наиболее надежных инструментов, которые могут помочь в обнаружении РИИ, что вероятно повлияет на решения о проведении тромболизиса и клинический исход [4, 116].

В исследовании Dippel D. W. J. и соавт. [108] было высказано предположение, что уровень правильной верификации КТ-признаков ишемического инсульта является важным показателем его исхода и эффекта тромболитической терапии. Авторы изучили эти признаки проспективно у 260 пациентов с ишемическим инсультом в бассейнах переднего кровообращения. Наличие признаков РИИ подтверждали путем динамического наблюдения пациентов в виде проведения КТ-исследований через 1 и 12 недель после поступления пациента. Рецензенты были опытными клиницистами, специализирующимися на лечении и диагностике инсульта (неврологи, нейрорадиологи и врачи эндоваскулярной хирургии). Среднее время между началом инсульта и выполнением первоначального КТ-исследования составило 3,2 часа. Более чем у половины пациентов (52%) при контрольной КТ был выявлен обширный инфаркт в бассейне средней мозговой артерии. Отношение шансов (хи - статистика) между наблюдателями для любого раннего признака инсульта составило 0,27 (95% доверительный интервал: от 0,15 до 0,39). Согласие

относительно распространенности ишемических изменений в бассейне средней мозговой артерии было незначимым: 0,37 и 0,35 соответственно [4, 108].

В исследовании Grotta J. C. и соавт. [26] выявили уровень согласия между нейрорадиологами с большим опытом в выявлении инсульта и другими клиницистами в отношении наличия признаков РИИ. Семьдесят первоначальных КТ-исследований были оценены 16 экспертами, включая неврологов региональных сосудистых центров (специализирующихся на диагностике и лечении инсультов), неврологов без опыта работы с инсультами, других врачей неотложной помощи, а также нейрорадиологов ургентной медицины. Значения Каппа, как критерия межэкспертного согласия (95% доверительный интервал [ДИ]), варьировались от 0,20 (-0,20 - 0,61) (слабое согласие) до 0,41 (0,37 - 0,45) (умеренное согласие). Авторы пришли к выводу, что даже среди опытных клиницистов существуют низкие уровни согласия в отношении распознавания и количественной оценки признаков ранних ишемических изменений [4, 26].

Исходя из данных зарубежной литературы, распознавание признаков тромбоза магистральных артерий головного мозга и ишемии на ранних стадиях является сложной диагностической задачей. Существуют противоречивые данные относительно способности врачей лучевой диагностики и, в том числе, рентгенологов с небольшим опытом работы, правильно трактовать КТ-признаки ишемического инсульта. В условиях ежегодного обновления кадрового состава больниц молодыми специалистами, наличия периода их адаптации к специфике работы нейрорадиолога, в сочетании с нехваткой рентгенологов для оперативной интерпретации лучевых исследований (в т.ч. КТ) в субъектах Российской Федерации, прогнозируется закономерное повышение уровня вариативности интерпретации лучевых изображений между врачами. И, как следствие, снижение качества оказания медицинской помощи [4].

1.2. Шкала ASPECTS: проблемы межэкспертной вариативности в клинической практике

С целью сформировать единый подход к диагностике ишемического инсульта в 2000 г. была разработана ASPECTS (Alberta stroke program early CT score) - полуколичественная шкала, созданная для оценки ранних ишемических изменений в бассейне средней мозговой артерии по 10 – бальной системе. ASPECTS предполагает оценку десяти областей на КТ-срезах на стандартных уровнях (уровне базальных ядер и ростральных структур) (Рисунок 5): хвостатое ядро, островок, чечевицеобразное ядро, внутренняя капсула и шесть других корковых областей (M1-M6). Участки M1–M3 находятся на уровне базальных ганглиев, участки M4–M6 — на уровне желудочков непосредственно над базальными ганглиями, границей является головка хвостатого ядра [3, 90].

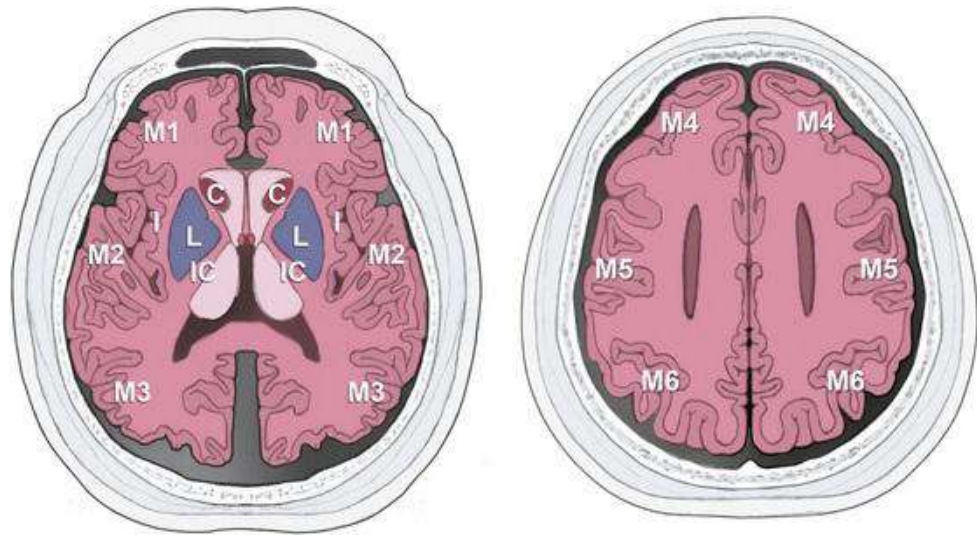


Рисунок 5 – Шкала ASPECTS (Alberta stroke program early CT score) [96]

Согласно шкале, последовательно оценивается каждый из десяти регионов. Подсчет по ASPECTS производится путем вычитания 1 балла из 10 для каждого достоверного признака ранних ишемических изменений (нарушение дифференцировки серого/белого вещества, сглаженность корковых борозд на стороне поражения, а также фокальная гиподенсивность) в каждом из регионов. При этом патологические изменения должны визуализироваться как минимум на

двух последовательных срезах для исключения эффекта объемного усреднения. Таким образом, КТ без патологических изменений соответствует 10 баллам по ASPECTS. Оценка в 0 баллов указывает на диффузное поражение всего бассейна средней мозговой артерии. Оценка ≤ 7 баллов коррелирует с худшим функциональным исходом и более высоким риском геморрагической трансформации инсульта [3, 57].

Шкала ASPECTS получила положительную оценку многих исследователей как надежный диагностический метод. К примеру, Baek J. H. и соавт. [91] в своей статье сделали вывод, что ASPECTS — лучший метод для прогнозирования функционального исхода у пациентов с острым ишемическим инсультом, получающих тромболитическую терапию, по сравнению с другими оценочными шкалами [3, 91]. Данный вывод подтверждается и другими исследованиями [111].

Однако использование шкалы имеет и ряд ограничений [69]:

- шкала ASPECTS используется в бассейне СМА (хотя на данный момент также известна модификация шкалы для оценки ишемии в вертебрально-базилярном бассейне (ВББ) — posterior circulation ASPECTS — pcASPECTS). Поэтому ее использование некорректно при окклюзии внутренней сонной артерии, в том числе при фетальном типе строения задней мозговой артерии [3, 106];
- оценка по ASPECTS затруднена на уровне M2-территории за счет артефактов от костей основания черепа [3];
- наличие субкортикальных и возрастных перивентрикулярных изменений белого вещества могут приводить к некорректной оценке по ASPECTS [3];
- изображения низкого качества, например, с двигательными артефактами, могут привести к ошибочной оценке по ASPECTS [3].

Значительным ограничением использования данной шкалы является уровень тромбоза средней мозговой артерии. Поскольку шкала охватывает всю территорию бассейна СМА, то ее использование корректно только при окклюзии проксимальных отделов артерии. Если сосуд тромбирован более дистально, тогда оценка всех 10-ти областей ASPECTS становится неточной (Рисунок 6).

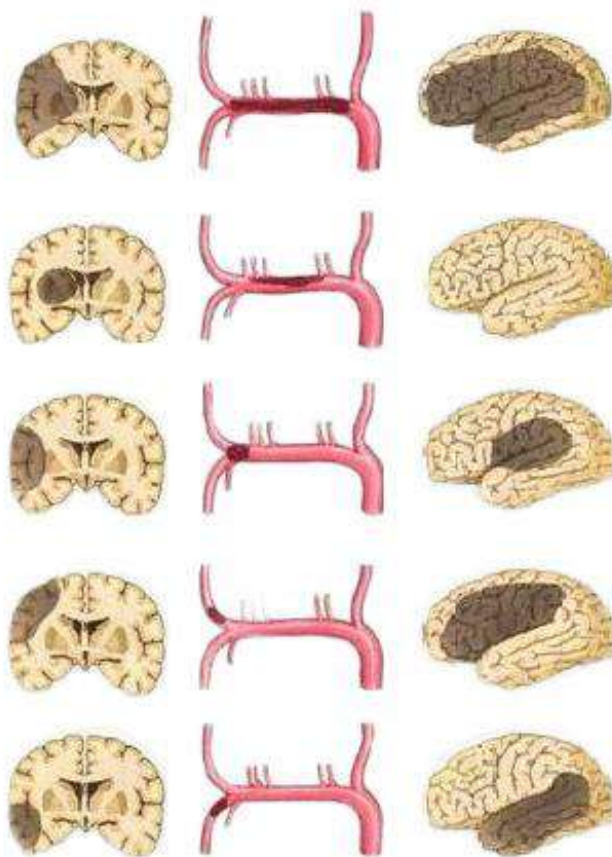


Рисунок 6 – Варианты локализации окклюзии средней мозговой артерии, ассоциированные с объемом поражаемой территории бассейна

Важно также отметить, что шкала не полностью стандартизирована. Области ASPECTS обозначены схематично. К примеру, некоторые затруднения вызывает оценка по шкале на «границе водоразделов». Расположение кортикальных пограничных зон может отличаться из — за развития лептоменингеальных коллатералей. В исходном исследовании ASPECTS для оценки использовались только два среза нативной КТ, но современной медицинской практике в подавляющем большинстве используются все доступные срезы сканирования [3, 115].

Еще одним источником межсубъектной вариативности являются характеристики РИИ, которые используются для подсчета баллов. Потеря дифференцировки коры, по данным некоторых авторов, связана с отеком и необратимым ее повреждением. В условиях эксперимента ишемия в виде

нарушение дифференцировки коры обратима только в течение нескольких минут после начала инсульта [3, 73]. В клинических условиях нарушение дифференцировки коры с фокальным снижением ее плотности с высокой специфичностью представляет собой необратимо поврежденную ткань мозга (то есть ядро инфаркта) [3, 54]. Поскольку на фоне отека в данных областях происходит сужение ликворных пространств, то сглаженность борозд является ассоциированным признаком нарушения дифференцировки коры. Однако в связи с недавними патофизиологическими исследованиями выявлена возможность визуализации изолированного сглаживания борозд (без нарушения дифференцировки), которое не считается достоверным признаком ишемии, учитываемым в ASPECTS, поскольку представляет собой потенциально обратимое явление [3, 50, 58].

Дополнительная критика ASPECTS заключается в том, что некоторые области — например, внутренняя капсула — по объему намного меньше других областей, но при этом данная структура имеет одинаковый «вес» в распределении баллов; таким образом, два пациента с одинаковыми баллами по ASPECTS могут иметь разную степень выраженности и объема РИИ [3, 101]. К примеру, ASPECTS подверглась критике в исследовании Phan T. G. и соавт [107]. Авторы обратили внимание, что в системе баллов ASPECTS стриатокapsулярная область (скорлупа, дорсолатеральные отделы хвостатого ядра, переднее бедро внутренней капсулы) имеет непропорциональный «вес». Четыре опытных нейрорадиолога оценили отдельные области ASPECTS на КТ-изображениях девятнадцати пациентов с инсультом в бассейне СМА со сравнением конечного объема ишемических изменений на КТ-исследованиях в динамике (на 5–10-й день). Объем ишемических изменений определяли путем ручной сегментации на КТ-изображениях. Для оценки регионального объема, связанного с каждой областью ASPECTS, использовалась линейная регрессия. Области ASPECTS имели не равный вес, при этом на стриатокapsулярную область приходилось 21% объема

РИИ в бассейне СМА при тотальном поражении территории кровоснабжения [3, 107].

Самый частый аспект критики данной шкалы заключается в том, что ASPECTS была введена с целью уменьшить вариативность межэкспертных оценок, однако исследования показали достаточно разнородные результаты уровня согласия между специалистами [3].

В исследовании Farzin В. и соавт. [52] был изучен уровень межэкспертного и внутриэкспертного согласия пятнадцати специалистов при оценке КТ-исследований по шкале ASPECTS у тридцати пациентов с ишемическим инсультом. В группу экспертов входили: шесть неврологов, специализирующихся на диагностике и лечении пациентов с ОНМК (четыре с опытом работы более десяти лет), пять рентгенэндоваскулярных хирургов (трое с опытом работы более десяти лет) и четыре нейрорадиолога (двое с опытом работы более десяти лет). Каждый специалист прошел два «слепых» этапа оценки с интервалом не менее трех недель. Дополнительно экспертам были предоставлены следующие клинические данные: пол, возраст, наличие / отсутствие афазии и гемипареза, а также оценка по шкале NIHSS (National Institutes of Health Stroke Scale). Межэкспертное согласие было слабым или умеренным, без существенной разницы между специальностями (нейрорадиолог, врач по рентгенэндоваскулярным диагностике и лечению, невролог) и опытом работы. Надежность между наблюдателями определялась с использованием критерия взвешенной каппы (k Коэна). Коэффициент согласия расценивался как умеренный при значении ($k > 0,4-0,6$), существенный ($k > 0,6-0,8$) и как полное согласие ($k > 0,8-1,0$). Ни одно из значений каппа не достигло существенного уровня ($0,6$) среди всех экспертов. Даже при дихотомическом делении по шкале ASPECTS на две категории (0– 5/5–10 баллов), межэкспертная согласованность не достигла существенного уровня ($k > 0,561$), это значит, что как минимум 5 из 15 экспертов вынесут разное заключение в 15% случаев. Внутреннее (внутриэкспертное) согласие (оценка эксперта при первом и повторном просмотре) варьировалось от

0,599 до 0,943. Врачи имели одинаковый уровень согласия при повторном просмотре в 40% случаев [52]. Также исследователи проанализировали статьи, оценивающие межэкспертное согласие при использовании ASPECTS, опубликованные с 2000 по 2015 г. Методологии проанализированных исследований отличались друг от друга по нескольким характеристикам, в том числе были ли врачам представлены клинические данные при оценке исследований, наличествовал ли лимит времени, доступ ко всем КТ-срезам исследований, возможность для экспертов устанавливать свои собственные параметры окна. Авторы сделали общий вывод, что данные факторы имеют весомое место в выявленной субъективности экспертной оценки, что отражает высокую степень вариабельности межэкспертного согласия [3, 52].

В исследовании Kobkitsuksakul С. и соавт. [75] был изучен уровень межэкспертного и внутриэкспертного согласия при оценке КТ-исследований сорока трех пациентов с ишемическим инсультом по шкале ASPECTS двумя нейрорадиологами (с опытом работы более десяти лет), стипендиатом в области нейрорадиологии (зарубежный вариант одной из ступеней последипломного образования) и ординатором. Согласие между двумя нейрорадиологами и стипендиатом в отношении дихотомического деления баллов по ASPECTS, оцененное с помощью каппа Коэна в большей степени являлось достаточно умеренным (0,486 - 0,678). Согласованность между двумя нейрорадиологами и ординатором по общему количеству баллов ASPECTS была в диапазоне от слабого до умеренного (0,198 - 0,491) [3, 75].

Coutts S. В. и соавт. [72] пришли к выводу, что оценка по ASPECTS варьируется в условиях анализа КТ изображений, выполненном проспективно и ретроспективно. При более высоких оценках ASPECTS (>7) специалист, проводящий исследование проспективно, имел тенденцию занижать баллы ишемических изменений. При более низких ASPECTS (<3) эксперт в настоящем времени напротив демонстрировал тенденцию преувеличивать объем ишемических изменений почти на 2 балла [3, 72]. Причины этого, вероятно,

отражают совокупность факторов, в том числе склонность системы визуального восприятия человека переоценивать границы областей поражения [3, 41]. Следует отметить, что Coutts S. B. и соавт. также указали на влияние человеческого фактора (желание не проводить тромболизис у пациентов с сопутствующей патологией и отягощенным анамнезом) при присвоении пациенту более низких баллов по шкале ASPECTS [3, 72].

В исследовании Alexander L. D. и соавт. [78], в которое включили пятьдесят пять пациентов с установленным диагнозом ишемического поражения головного мозга в подострой стадии, три невролога ретроспективно оценили с использованием ASPECTS КТ-изображения, полученные через два дня и более от начала заболевания. Авторы отмечают, что согласие между специалистами было практически полным со значением k Коэна равной 0,90 [3, 78].

Van Seeters T. и соавт. [97] провели проспективное исследование с включением в него ста пяти пациентов с острым неврологическим дефицитом и подозрением на острый ишемический инсульт в расширенном (9 часов после появления симптомов) “инсультном окне”. Всем пациентам при поступлении были выполнены бесконтрастное КТ - исследование головного мозга, КТ-перфузия и КТ-ангиография сосудов головного мозга. Все изображения оценивались дважды по двум параметрам: наличие РИИ с их оценкой по шкале ASPECTS, а также с помощью правила «одной трети» бассейна средней мозговой артерии (оценочная шкала распространенности ишемических изменений, разделяющая бассейн СМА на три части, при пораженности более $\frac{1}{3}$ территории тромболизис не проводился). Четыре нейрорадиолога дважды оценили КТ-изображения данных пациентов на предмет меж- и внутриэкспертного согласия с использованием каппа статистики и коэффициента внутриклассовой корреляции. В результате, согласие между экспертами по правилу $\frac{1}{3}$ СМА и ASPECTS варьировалось от умеренного до существенного при оценке бесконтрастных КТ, от низкого до существенного при оценке изображений КТ-ангиографии, и стало полным для всех карт КТ- перфузии. Исследователи сделали вывод, что КТ-

перфузия является более надежным методом выявления ишемического инсульта с помощью компьютерной томографии с точки зрения уровня межэкспертного согласия [3, 97].

Рехман J. Н. и соавт. [111] опросили шесть врачей на предмет их трактовки ASPECTS. Половина исследователей включали в оценку изолированное набухание коры (без нарушения ее дифференцировки). Несмотря на это, уровень согласия между специалистами в отношении ASPECTS был выше, чем межэкспертная согласованность относительно правила 1/3 СМА [3, 111].

В своем исследовании Mc Taggart R. А. и соавт. [27] сравнили уровень межэкспертной вариативности в оценке КТ- и МР- ASPECTS на ранних стадиях инсульта. Проспективно проанализированы семьдесят четыре пациента, которым выполнялись КТ и МРТ, а также была проведена тромбоэкстракция в течение 12 часов от начала инсульта. Два эксперта оценили наличие и объем ранних ишемических изменений на бесконтрастных КТ- изображениях с помощью ASPECTS, а также на диффузионно- взвешенных изображениях (ДВИ). Согласие между экспертами в отношении КТ- ASPECTS и ДВИ- ASPECTS составило 0,579 и 0,867 соответственно. ДВИ-ASPECTS коррелировал с функциональным результатом ($p=0,004$), тогда как КТ- ASPECTS — нет ($p=0,534$). И ДВИ- ASPECTS, и КТ- ASPECTS коррелировали с объемом ишемии, выявляемой на ДВИ. Анализ рабочих характеристик показал, что ДВИ-ASPECTS превосходит КТ- ASPECTS в соотношении величины временного интервала между появлением симптомов и прогнозированием хорошего функционального результата. По результатам данного исследования установлено, что согласие между экспертами при оценке по ДВИ- ASPECTS было выше, чем при оценке по КТ- ASPECTS. Методика оценки РИИ с помощью ДВИ- ASPECTS также превзошла КТ- ASPECTS в отношении прогнозирования функционального результата (90 дней) [3, 27].

В исследовании Varber P. А. и соавт. [115] были проанализированы двести три КТ-исследования для установления взаимосвязи между количественной

оценкой баллов по шкале на КТ изображениях и прогнозированием исхода ОНМК при проведении тромболитической терапии. Исходное значение ASPECTS обратно коррелировало с тяжестью инсульта по шкале NIHSS ($r=-0,56$, $p < 0,001$). ASPECTS предсказывала функциональный исход и появление геморрагической трансформации инсульта ($p < 0,001$, $p=0,012$ соответственно). Чувствительность ASPECTS к функциональному результату составила 0,78, а специфичность — 0,96. Согласие между экспертами в отношении ASPECTS было достаточно высоким (каппа 0,71–0,89). Исследователи в своей статье сделали вывод, что шкала ASPECTS является надежным методом прогнозирования функционального исхода после тромболитической терапии [3, 115]. Собственно, введение шкалы как оценочной методики в диагностику ишемического инсульта согласовывался, по итогам этого исследования, с началом международного использования тромболитической терапии [3].

Еще одной причиной вариабельности оценки между экспертами по данной шкале может быть разнородность групп специалистов, участвующих в исследованиях [3]. Wilson A. T. и соавт. [81], указывают, что опыт, уровень подготовки и медицинская специализация также могут влиять на оценку. В большинстве исследований, проверяющих надежность и межэкспертное согласие, принимали участие опытные неврологи и/или нейрорадиологи; в ограниченном числе статей упоминаются ординаторы и начинающие специалисты [3, 21, 45]. Однако даже среди опытных врачей могут существовать расхождения в трактовке ASPECTS в разных странах [69]. Другие важные, но не решаемые причины изменчивости оценки ASPECTS включают следующие факторы: программное и аппаратное обеспечение, местонахождение эксперта, время суток, параметры срочности выдачи заключения (цейтнот), а также его личностные качества [3, 69].

Несмотря на то, что шкала ASPECTS — это метод, который широко применяется в современной клинической практике, ее основным недостатком является возможность вариативности экспертных оценок, в том числе среди врачей с малым стажем, поскольку высокие показатели межэкспертного согласия

достигались в основном в группах с опытными рентгенологами. Выраженное многообразие результатов и низкий уровень межэкспертного согласия не позволяют в настоящее время считать данную шкалу действительно надежным вариантом стандартизированной оценки и может отражаться на дальнейшем лечебном процессе [3, 5].

В этой связи перспективным выглядит внедрение в клиническую практику методов полуавтоматической и автоматической обработки КТ-изображений с применением систем искусственного интеллекта, которые в перспективе могут улучшить стандартизацию оценки.

1.3. Возможности систем автоматического анализа компьютерно- томографических изображений в диагностике ишемического инсульта в бассейне средней мозговой артерии, а также в применении шкалы ASPECTS

Искусственный интеллект — это комплекс технологических решений, имитирующих когнитивные функции человека (включая самообучение и поиск решений без заранее заданного алгоритма), позволяющий при выполнении задач достигать результаты, как минимум сопоставимые с интеллектуальной деятельностью человека. Это понятие включает информационно-коммуникационную инфраструктуру и программное обеспечение, в которых используются методы машинного, репрезентативного и глубокого обучения, а также процессы обработки данных и выработки решений [6, 9]. ИИ — одна из быстро развивающихся областей информатики и вычислительной техники, имеющая большое значение для радиологии. Наиболее разработанным и используемым классом методов искусственного интеллекта является машинное обучение. Оно используется для частичной или полной автоматизации решения сложных профессиональных задач на основе накопленных данных. Алгоритмы развиваются при увеличении объемов доступных баз данных, улучшают свою результативность с опытом, а также учатся давать конкретные ответы, обрабатывая большие объемы информации [6, 100]. Перед оценкой моделей

машинного обучения необходима спецификация медицинских диагностических задач, поскольку модели должны быть соответствующим образом обучены. Это достигается с помощью основных типовых задач: обучение с учителем (supervised learning), обучение без учителя (unsupervised learning) и частичное обучение (semisupervised learning) [6]. Эти типы задач обучения по прецедентам являются фундаментальными стратегиями, применяемыми в зависимости от имеющихся данных. При обучении с учителем искусственный интеллект извлекает информацию из определенной части обученных выборок с верифицированной и размеченной патологией для того, чтобы предсказывать результаты неизвестных данных [6, 44]. И наоборот, при обучении без учителя происходит сравнение «нормы» с базами данных неразмеченной патологии для реализации задачи кластеризации, заключающейся в том, чтобы сгруппировать объекты в кластеры, используя данные о попарном сходстве объектов. При частичном обучении происходит комбинация двух вышеуказанных методов и требуется только небольшая часть маркированных обучающих данных. Неразмеченные изображения также используются при обучении [6, 102]. Необходимо, чтобы размеченные данные были достоверны. Именно поэтому была введена концепция «основа истины», которая определяет проверку результатов машинного обучения на предмет точности и является фундаментальной для проверки эффективности программ ИИ. «Основой истины» можно считать данные, которые являются достоверно верифицированными (подтвержденные «золотым стандартом») [6, 43]. Глубокое обучение является совокупностью методов машинного обучения и ведущим направлением большинства инструментов ИИ для интерпретации изображений. Глубокое обучение — это алгоритмы, представленные в виде нескольких уровней обработки информации, связанных между собой. Эта многослойная система нелинейных фильтров используется для извлечения признаков с преобразованиями, что подразумевает под собой получение каждым последующим слоем на входе выходных данных предыдущего слоя. Такие многослойные алгоритмы образуют большие искусственные нейронные сети [6,

102]. Искусственные нейронные сети должны быть «обучены» с использованием стандартизированных наборов данных. В лучевой диагностике они обычно состоят (по крайней мере, на начальном этапе) из маркированных вручную изображений, используемых алгоритмами для обучения (сегментация). После того как сеть будет обучена, ее необходимо протестировать с использованием другого набора данных, предназначенного для оценки соответствия модели обучения требуемому результату [6]. На этом этапе часто наблюдается «переоснащение» модели. Yamashita R. и соавт. [42] описывают переоснащение как ситуацию, «когда модель изучает статистические закономерности, характерные для обучающего набора, то есть в конечном итоге запоминает нерелевантный шум вместо изучения данных и, следовательно, работает хуже на последующем новом наборе данных. Следствием переоснащения является то, что сеть не будет обобщать данные, с которыми ранее не встречалась, и станет совершать большее количество ошибок [6, 42]. Между объемом, а также качеством первичных данных и соответствием модели требуемому результату существует прямая корреляция. Поскольку ввиду возможности проведения обучения в несколько этапов и тестирования на разнородных наборах данных постепенно улучшается производительность сети, то прежде, чем алгоритм будет выпущен для общего использования, оцениваются его точность и универсальность. Другим решением является так называемое «дополнение данных», которое означает изменение обучающих данных путем добавления некоторой новой информации [6, 42]. Как правило, чем «глубже» сеть (больше уровней) и чем больше циклов обучения, тем выше производительность сети. Основополагающим направлением применения ИИ в нейрорадиологии является автоматическая сегментация поражений головного мозга, позволяющая разгрузить врача рентгенолога от трудоемкой функции выполнения сегментации вручную [6, 29]. При клиническом применении рентгенологи в основном выполняют ручную сегментацию, которая является субъективной, занимает большое количество рабочего времени и плохо воспроизводима при повторной

процедуре. Автоматическая сегментация полностью контролируется алгоритмом без участия человека. Скорость сегментации высокая, а результаты воспроизводимы [6, 25]. Другим применением ИИ в радиологии является радиомика. Это способность представления сложных объектов в виде совокупности количественных признаков [6].

Программные продукты на основе радиомики позволяют извлекать признаки из диагностических изображений, конечным продуктом которых является определение параметра конкретной патологии [6]. Радиологический анализ может получить большое количество элементов из области, представляющей интерес в исследовании КТ или МРТ, и сопоставить эти характеристики друг с другом и другими данными [61]. Системы машинного обучения в настоящее время применяются для диагностики и классификации опухолей головного мозга [70], определенных психических расстройств [37], эпилепсии [24], нейродегенеративных [99] и демиелинизирующих расстройств [84]. Алгоритмы машинного обучения были также разработаны и для помощи в диагностике и индивидуальном выборе лечения при остром ишемическом инсульте. Одним из наиболее важных клинических критериев для успешного проведения тромболиза с применением тканевого активатора плазминогена при остром ишемическом инсульте является его проведение в первые 4,5 часа от появления симптомов, но обычно начало инсульта неизвестно [6]. Для решения этой проблемы Но К. S. и соавт. [39] разработали алгоритм глубокого обучения, основанный на архитектуре автокодера, с извлечением признаков визуализации из изображений МР-перфузии (PWI) и определения времени, прошедшего с момента начала инсульта. Chen L. и соавт. [38] на основе данных семисот сорока одного пациента и модели глубокого обучения, включавшей две нейронные сети, разработали алгоритм сегментации поражений головного при инсульте с помощью DWI-изображений. Измерение перфузионно-диффузионного несоответствия и расчет вероятности формирования зоны ядра инфаркта с использованием подходов на основе МРТ для оценки тканей, подверженных

риску, могут применяться при принятии решений о типе лечения инсульта [6]. Bouts M. J. и соавт. [53] проанализировали способность пяти алгоритмов выявлять потенциально жизнеспособные ткани мозга на основе МР-изображений крыс, подвергшихся правосторонней окклюзии средней мозговой артерии (СМА) без последующей реперфузии, со спонтанной реперфузией или реперфузией, вызванной тромболизисом. Наивысшая точность идентификации ишемической ткани, которая может восстановиться после последующей реперфузии, наблюдалась с использованием обобщенной линейной модели (коэффициент сходства Дайса = $0,79 \pm 0,14$). Аналогичным образом Huang S. и соавт. [67] применили метод опорных векторов (SVM) для прогнозирования инфаркта у крыс на попиксельной основе с использованием показателей скорости мозгового кровотока (CBF) и изображений измеряемого коэффициента диффузии (ADC) по данным МРТ. Еще одним применением систем машинного обучения при остром ишемическом инсульте является прогнозирование факторов, которые будут способствовать ухудшению неврологического статуса и повышения вероятности развития отека головного мозга [6]. Chen Y. и соавт. [35] предложили определения и измерения объема спинномозговой жидкости в динамике, поскольку он может представлять собой чувствительный биомаркер развития и прогрессирования отека мозга. Первоначальная когорта состояла из пятидесяти пяти результатов КТ-обследования. Предварительная обработка проводилась с использованием модели обобщенных оценочных уравнений для расчета объема спинномозговой жидкости в головном мозге с поправкой на возраст. Результаты исследования показали, что уменьшение объема спинномозговой жидкости в динамике коррелировало с объемом инфаркта, наличием отека головного мозга и степенью смещения срединных структур [6]. Для сравнения Dhar R. и соавт. [30] представили автоматизированный метод сегментации ликворных пространств с помощью ансамбля машинного обучения на основе метода случайных лесов с сегментацией геометрического активного контура. У 38 пациентов была произведена разметка ликворных пространств

головного мозга в течение 6 часов от начала инсульта, а затем через 24 часа от дебюта заболевания. Такой подход позволил точно отследить сужения субарахноидальных пространств головного мозга как корреляционный показатель объема СМЖ. Коэффициенты корреляции Пирсона между изменениями объема СМЖ и нормальными показателями оказались статистически значимыми. Разработанные алгоритмы представляют собой потенциал для будущих исследований и могут служить биомаркером тяжести отека мозга. Исход пациентов с острым ишемическим инсультом зависит от качества и своевременности лечения, поэтому следует учитывать риски осложнений при принятии решения о выборе определенного типа терапии. Yu Y. и соавт. [92] разработали метод прогнозирования локализации и степени геморрагической трансформации при инсульте, как наиболее тяжелого осложнения после реперфузионной терапии. Перфузионные и DWI-изображения ста шестидесяти пяти пациентов, получавших реперфузионную терапию, были проанализированы с использованием пяти подходов машинного обучения, при этом спектральная регрессия ядра продемонстрировала точность $83,7 \pm 2,6\%$. В многоцентровом ретроспективном исследовании [6, 82] ученые оценили прогностическую способность геморрагической трансформации по данным МР-перфузии. Были собраны данные МР-перфузионных изображений у 263 пациентов из четырех медицинских центров, которые служили входными данными для линейных и нелинейных прогностических моделей со средней точностью $>85\%$ при прогнозировании ИИ [6]. Nielsen A. и соавт. [94] провели исследование на основе созданной ими нейронной сети с 9 биомаркерами в качестве входных данных для расчета объема поражения у пациентов, получавших внутривенную тромболитическую терапию (ВТТ). Сравнивались исходные данные 35 пациентов, получавших внутривенно ВТТ и 29 пациентов из контрольной группы. Эта модель предсказала с точностью 88% конечный объем инфаркта. Bentley P. и соавт. [93] на основе данных ста шестнадцати пациентов (результаты компьютерной томографии и клиническая оценка тяжести), получавших ВТТ,

разработали систему, позволяющую предсказать риск симптоматического внутримозгового кровоизлияния после внутривенной терапии тромболизисом [6]. В их исследовании система на основе метода опорных методов давала лучший прогноз по сравнению с традиционными инструментами прогнозирования, основанными на данных экспертной оценки, такими как внутримозговое кровоизлияние после тромболизиса, ранние признаки инфаркта, признак гиперденсной мозговой артерии, возраст и баллы по шкале NIHSS (SEDAN) [6].

Алгоритмы машинного обучения могут помочь и в прогнозировании двигательных нарушений у пациентов с инсультом. Forkert N. D. и соавт. [83] при обработке МРТ-изображений применили 12 классификационных моделей на основе метода опорных векторов для расчета соответствующей оценки пациентов с ишемическим инсультом по модифицированной шкале Рэнкина (Modified Rankin scale — mRS) в течении 30 дней — с помощью параметров, включающих вовлечение в поражение разных областей мозга, латеральность инсульта и другие дополнительные характеристики, такие как объем инфаркта, NIHSS при поступлении и возраст пациента. Превосходное прогнозирование неврологических нарушений по mRS наблюдалось за счет интеграции дополнительных функций и предоставления информации о локализации, с точностью прогнозирования mRS по нескольким значениям 56% и дихотомической точностью прогнозирования mRS (0–2 против 3–5) 85% [6]. В исследовании Rondina J. M. и соавт. [48] была создана модель для прогнозирования двигательного дефицита верхних конечностей у 50 пациентов с инсультом, разработанная на основе данных структурной МРТ вместо функциональной. Таким образом, в настоящее время существуют успешные попытки применить все возможности систем искусственного интеллекта для оценки данных нейровизуализации при ОНМК: раннее выявление диагностическими методами визуализации, оценка времени начала заболевания, сегментация поражения, анализ наличия и возможности возникновения отека мозга, а также прогнозирование осложнений и результатов лечения. Однако

существует ряд ограничений для дальнейшего развития систем искусственного интеллекта [6]. Первое ограничение — это размер выборки. Алгоритмы глубокого обучения с использованием медицинской визуализации часто требуют существенный массив данных, который, в силу своей специфики, может быть недоступен. Например, алгоритм машинного обучения, продемонстрировавший превосходную производительность при дифференциальной диагностике между злокачественными и доброкачественными новообразованиями кожи, в сравнении с результатами экспертной оценки 21 дерматолога, был обучен на наборе данных из почти 130 000 изображений [6, 49]. Набор данных такого размера для публичного использования, в настоящее время не существует. Препятствия в обмене данными между учреждениями, а также отсутствие финансирования для правильной предварительной обработки и курирования этих изображений и ограничения на размещение такого набора данных являются причиной некоторых задержек в создании этого репозитория [6].

Еще одно ограничение, встречающееся в методах машинного обучения на основе нейровизуализации — это необходимость маркировки областей интереса или результатов «золотого стандарта» на изображениях. Другими словами, помимо сбора изображений, необходима их разметка, идентификация и сегментация. Например, для обучения алгоритма, оценивающего наличие или отсутствие гиперденсной СМА, необходима предварительная разметка «учителем». С учетом того, что человеческий экспертный ресурс является достаточно ограниченным, есть запрос на снижение необходимости его применения в обучении моделей искусственного интеллекта. При этом следует отметить, что большинство результатов представленных авторами алгоритмов в настоящее время не прошли независимую оценку в клинической практике. Медицинскому сообществу предстоит широкая клиническая апробация разработанных систем искусственного интеллекта на независимых наборах данных. Требуется проведение рандомизированных исследований, оценивающих отдаленные результаты систем искусственного интеллекта, что может внести

изменения в показатели диагностической точности данных алгоритмов. Также на основе независимой клинической апробации возможно рекомендовать конкретные программные продукты в медицинскую практику. Отдельным пунктом является изучение взаимодействия врача и систем машинного обучения, а также влияние этой коллаборации на процесс принятия решений, показатели качества медицинской помощи, продолжительность и качество жизни пациентов [6].

Для частичного решения проблем субъективности оценки по шкале ASPECTS рядом авторов предлагается применение систем искусственного интеллекта. На сегодняшний день программы автоматического анализа изображений сосредоточены на анализе бесконтрастных КТ, КТ-ангиографии и визуализации перфузии на основе КТ или МРТ. Они направлены на определение и количественную оценку ядра инсульта, пенумбры, статуса коллатерального кровотока и локализации артериальной окклюзии в автоматическом режиме. Один из вариантов программного обеспечения для диагностики ишемических изменений представили Wolff L. и соавт. [113] сравнили его эффективность с аналитическими способностями врачей-специалистов. В их исследовании была получена специфичность на уровне 89–89%, чувствительность 41–57% и точность 0,750–0,795. Авторы сделали вывод, что диагностическая точность данной системы сравнима с показателями врачей, участвующих в исследовании и может помочь рентгенологом в обнаружении ранних ишемических изменений [6, 113].

Также ряд исследований сосредоточены на том, чтобы улучшить межэкспертную согласованность посредством внедрения систем автоматического анализа. Так, Delio P. R. и соавт. установили, что применение алгоритмов искусственного интеллекта увеличивает согласованность экспертов от 72 до 78% [6, 34].

Также, Culbertson C. J. и соавт. сделали вывод, что с помощью автоматической сегментации возможно увеличить межэкспертную

согласованность среди экспертов с небольшим опытом диагностики ишемического инсульта [6, 109].

В настоящий момент проходят апробацию несколько клинических программных продуктов для помощи рентгенологам в интерпретации изображений при остром ишемическом инсульте, способные автоматически оценивать данные КТ с присвоением балла по шкале ASPECTS. Ограничением данных исследований является небольшое число данных, включенных в исследование. Также в исследованиях не анализируется влияние различных алгоритмов реконструкции на диагностические показатели эффективности текущих вариантов систем. Но первые результаты показывают, что применение данных алгоритмов позволяет улучшить межэкспертное согласие при оценке по шкале ASPECTS [6, 34, 40]. В то же время алгоритмы систем автоматического анализа не предназначены для использования в качестве автономных диагностических инструментов. Они могут помочь врачам получить более точную и стандартизированную интерпретацию результатов КТ- и МРТ-исследований, что может улучшить тактику лечения пациентов и функциональный исход заболевания [6, 31].

В этой связи перспективным выглядит внедрение в клиническую практику методов полуавтоматической и автоматической обработки КТ-изображений с применением систем искусственного интеллекта, которые по результатам первых исследований улучшают стандартизацию оценки. Но для полноценного принятия таких систем в клиническую практику необходима их клиническая апробация на независимых наборах разных данных [6, 33].

Согласно Obuchowski N. A. и соавт. [88] существует четыре варианта использования алгоритмов ИИ в медицинской визуализации: (1) модель первого или параллельного чтения, когда алгоритм ИИ сначала оценивает изображение, а затем результаты этой оценки просматриваются человеком, (2) модель второго чтения, когда алгоритм ИИ применяется после того, как человек интерпретировал данные, для возможности включения дополнительных результатов,

обнаруженных ИИ, (3) модель сортировки, когда алгоритм распределяет случаи в соответствии с наличием (в процентном отношении) подозреваемой патологии (4) модель предварительного скрининга, когда алгоритм ИИ применяется к набору изображений для выявления нормы с генерированием клинического отчета, и дальнейшей оценкой оставшихся (классифицируемых как “не норма”) случаев врачом (Рисунок 7).



Рисунок 7 – Модели использования алгоритмов искусственного интеллекта в медицинской визуализации [88]

Данные модели могут быть использованы только при правильной диагностике патологических изменений алгоритмами искусственного интеллекта. В этом случае программами должны быть достигнуты и доказаны следующие параметры: высокая точность, специфичность и чувствительность. Однако несмотря на определенный Obuchowski N. A. и соавт. [88] порядок взаимодействия врача и моделей искусственного интеллекта, все же остается неясным, как отразятся данные виды комплементарной оценки на диагностической эффективности врачей, какая модель является наиболее результативной и не обусловит ли подобная коллаборация отрицательную динамику правильных интерпретаций лучевых изображений.

Также, существует ряд ограничений для внедрения алгоритмов искусственного интеллекта. К ним относятся не только практические соображения, такие как совместимость данных технологий с другими системами, включая PACS и электронные медицинские карты (истории болезни), но и необходимость проверки точности программ, поскольку большинство результатов представленных алгоритмов в настоящее время не прошли независимую оценку в клинической практике. Из-за растущей сложности и разнообразия инструментов искусственного интеллекта для нейрорадиолога не всегда очевидны технические характеристики этих приложений, особенно потому, что разработчики очень редко предоставляют полную информацию о своих продуктах. Таким образом, врач сталкивается с дилеммой: основывать или нет свое решение на результатах алгоритмов искусственного интеллекта, не зная подробно, что происходит внутри «черного ящика» этих систем. Эта дилемма усугубляется тем, что точность алгоритмов напрямую зависит от характеристик, качества и структуры баз данных (томограф, толщина среза, наличие артефактов технической и динамической нерезкости), используемых для обучения, а также квалификации специалистов, выступающих «учителями» программного продукта. Зачастую, при обучении системы автоматического анализа используются базы данных, находящиеся в открытом доступе, без надлежащей верификации патологии, что

не может повысить уровень доверия врачей к алгоритмам искусственного интеллекта.

Крайне важным фактором внедрения систем искусственного интеллекта является эргономичность каждой отдельной системы автоматического анализа. При данной интеграции должно быть доказано, что их эксплуатация занимает не только малое количество времени, но и удобна в использовании. В том числе, критически важным аспектом остается изменение регламента рабочего процесса врача-рентгенолога.

В то же время, отсутствие структурированного подхода к внедрению технологий искусственного интеллекта и значительные различия в способах использования подобных систем [32] не может не отразиться на качестве оказания медицинской помощи, функциональном исходе заболевания и требует детализированного изучения с предоставлением наиболее эффективной модели комплементарной оценки различных патологических изменений.

Таким образом, только в случае доказанной во всех аспектах правомочности и эффективности внедрения систем автоматического анализа с улучшением условий работы врача рентгенолога и повышением квалифицированности его оценки, имеет место дальнейшее обсуждение успешной интеграции искусственного интеллекта в здравоохранение.

ГЛАВА 2. ОБЩАЯ ХАРАКТЕРИСТИКА КЛИНИЧЕСКОГО МАТЕРИАЛА И МЕТОДЫ ИССЛЕДОВАНИЯ

2.1. Разработка базы данных компьютерно- томографических исследований пациентов с клинической картиной острого нарушения мозгового кровообращения в бассейне средней мозговой артерии

Для изучения качества интерпретации компьютерно- томографических томограмм врачами-рентгенологами, а также диагностической эффективности систем автоматического анализа на втором этапе диссертационного исследования была разработана база данных [19], состоящая из бесконтрастных КТ-исследований головного мозга 150 пациентов с клинической картиной ОНМК в бассейне средней мозговой артерии, поступивших в региональный сосудистый центр г. Санкт- Петербурга в период с 1 декабря 2020 г. по 30 декабря 2021 г. с подозрением на острый ишемический инсульт (согласно шкале догоспитальной оценки тяжести инсульта Лос-Анджелесской шкалы моторного дефицита - LAMS).

Клиническая картина всех пациентов включала синдром средней мозговой артерии (контралатеральные очагу поражения гемиплегия или гемипарез, гемигипестезия, гемианопсия). У 43 пациентов (35 %) гемипарез в большей степени был выражен в верхних конечностях. У 15 пациентов (1%) пациентов выявлялся парез взора в сторону поражения. У 84 пациентов (56%) были диагностированы различные виды афазий – эфферентная и афферентная моторные афазии, сенсорная афазия, их сочетание [12]. У 10 пациентов (6%) клиническая картина включала грубую очаговую полушарная симптоматика по гемитипу, в том числе тотальную афазию или анозогнозию. У пациентов с исключенным ишемическим инсультом в последующем были установлены парез Тодда, гипертензионная энцефалопатия, психогенные расстройства или же гипогликемия.

Всем пациентам было проведено комплексное клинико-неврологическое исследование, включавшее тщательный сбор и анализ жалоб, анамнеза с учетом сопутствующей соматической патологии, объективный и неврологический осмотр. Больные обследовались по стандартной схеме исследования неврологического пациента. Степень неврологического дефицита и тяжести инсульта оценивалась при поступлении и в динамике по шкале NIHSS, активность в повседневной жизни – при помощи индекса Barthel и Rivermid, интеллектуально-мнестические нарушения – по шкале MMSE (Mini-mental State Examination), функциональная дееспособность – по шкале Рэнкина в день поступления и при выписке из стационара.

Сопутствующая соматическая патология была диагностирована у 120 пациентов (80% пациентов): артериальная гипертензия у 101 пациента (67%), сахарный диабет у 23 пациентов (15%), мерцательная аритмия у 20 пациентов (13%), инфаркт миокарда у 3 пациента (2%).

У 100 пациентов ишемический инсульт в бассейне средней мозговой артерии был подтвержден, у 50 пациентов исключен. У 79 (79%) пациентов был диагностирован кардиоэмболический подтип ишемического инсульта, у 21 (21 %) – атеротромботический, согласно классификации TOAST (Trial of Org 10172 in Acute Stroke Treatment). Данным пациентам диагноз ишемического инсульта был установлен врачом неврологом в соответствии с рекомендациями МЗ РФ и верифицирован с помощью КТ- ангиографии брахиоцефальных и церебральных артерий, а также данных КТ- перфузии [12]. Пациентам с исключенным ишемическим инсультом были проведены КТ- ангиография брахиоцефальных и церебральных артерий, а также КТ- перфузия и контрольное КТ исследование (через 24 часа), не выявивших патологических изменений.

Компьютерные томограммы были получены на компьютерном томографе GE Revolution EVO 128 (148 исследований) и Toshiba Aquilion 64 (2 исследования).

КТ- исследования проводились в стандартной укладке - лежа на спине, руки вдоль туловища, в головной подставке, без задержки дыхания. Сканирование проводилось от конвекса до уровня атлантаксиального сочленения. Размер матрицы составлял 512x512, толщина среза 5 мм, с возможностью реконструкции до 1,3 мм (алгоритмы реконструкции - convolution Kernel), питч-фактором 0,53, DFOV (поле обзора отображения) по оси X/Y был равен 180-324 мм.

Для подтверждения или исключения ОНМК у пациентов дополнительно были проведены КТ- ангиография брахиоцефальных и церебральных артерий, а также КТ перфузия.

КТ- ангиография брахиоцефальных и церебральных артерий также проводилась в стандартной (вышеописанной) укладке, с введением неионного контрастного препарата (Омнипак 300 мг/мл) в объеме 80 мл при помощи автоматического шприца-инъектора. При его использовании скорость введения составляла 2-3 мл/с с задержкой начала сканирования 50-60 с. Диапазон томографии – от дуги аорты до конвекса головного мозга. Введение контрастного препарата выполнялось по жизненным показаниям без учета общих противопоказаний (креатинин сыворотки >1,5 мг/дл (>130 мкм/л) и гипертиреоз). Размер матрицы составлял 512x512, толщина среза - 0,6 мм, DFOV (поле обзора отображения) по оси X/Y 180–412 мм.

КТ- перфузия проводилась с внутривенным контрастированием неионными контрастным препаратом (Омнипак) с концентрацией йода 300 мг\мл, объемом 50 мл, болюсно автоматическим шприцом-инъектором в кубитальную вену со скоростью 5-7 мл\секунду. Регистрация данных начиналась через 4 секунды после начала введения контрастного вещества и продолжалась в течение 50 секунд с интервалом в 1 секунду. Параметры томографии: коллимация срезов 4-8 мм, напряжение на трубке – 80 кВ, 200 мА, время оборота трубки – 1 с. Сканирование проводилось без задержки дыхания, блок срезов (8,5 мм) располагался с возможностью охвата наибольшей территории б. СМА. Постобработка данных динамической серии сканирования проводилась в специализированном

программном приложении. Эффективная доза облучения для всего комплекса КТ-исследований составила в среднем 20 мЗв.

Все данные были деперсонализированы с помощью программы Launch DicomAnonymize.

В базу данных, после верификации методом компьютерной ангиографии брахиоцефальных и церебральных артерий, а также КТ-перфузии, включены 100 пациентов с подтвержденным тромбозом средней мозговой артерии и признаками ишемических изменений в острой и острой стадиях в данном бассейне, 50 пациентов, у которых диагноз ОНМК в бассейне СМА не подтвердился (в том числе при контрольном бесконтрастном КТ-исследовании через 24 часа). Свидетельство о регистрации базы данных RU 2022620850 [19].

База данных объединяет: данные нативной КТ головного мозга (форматы DICOM), КТ-ангиографии сосудов головного мозга, брахиоцефальных артерий (формат DICOM), КТ-перфузии (наличие зон пенумбры и процентное соотношение ядра инсульта от данного значения, форматы DICOM и JPEG); сведения о патологическом процессе - наличии признаков острейшего и острого нарушения мозгового кровообращения, наличии изменений на контрольном КТ-исследовании (геморрагическая трансформация инсульта, отсутствие наличия сформированной ишемии) (Рисунок 8).

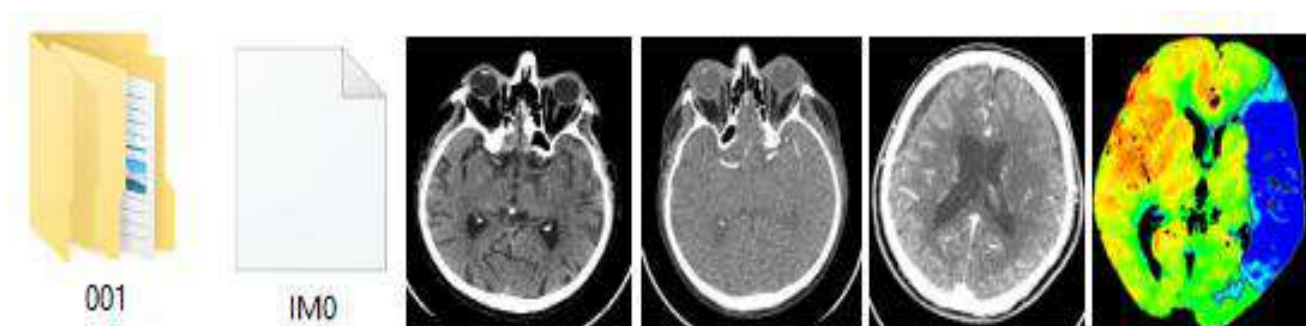


Рисунок 8 – Пример данных из базы компьютерно- томографических исследований пациентов с клинической картиной острого нарушения мозгового кровообращения в бассейне средней мозговой артерии

База данных включает также такую информацию, как пол и возраст пациентов (Таблица 1), срок выполнения КТ от начала заболевания (до/ более 6 часов), бассейн ОНМК (левая/правая средняя мозговая артерия), сегмент средней мозговой артерии, в котором был выявлен тромбоз (Таблица 2), особенности развития Виллизиева круга (задняя/передняя трифуркация ВСА/ замкнутость), наличие признаков ишемического инсульта (оценка производилась на основе консенсуса трех нейрорадиологов с опытом более 10 лет в ургентной медицине, а также на основе данных КТ- перфузии [46]), баллы по шкале ASPECTS (Таблица 3), данные КТ – перфузии (наличие пенумбры, процент соотношения ядра и пенумбры), данные о проведении тромбэкстракции, наличие контрольного бесконтрастного КТ- исследования (баллы по ASPECTS/ данные по Гейдельбергской классификации, критериям ECASS), исходе (Таблица 4), а также технические данные (томограф/наличие динамических артефактов/толщина среза (бесконтрастная КТ/АГМ)/ алгоритмы реконструкции (convolution Kernel)).

Таблица 1 – Распределение данных о поле и возрасте пациентов, включенных в базу данных

Группа пациентов	Количество	Средний возраст	Женский пол	%		Мужской пол	%	
				Женский	Мужской		Женский	Мужской
Пациенты с подтвержденным ОНМК	100	72,7 (36 - 95)	52	52	51,3	48	48	48,7
Пациенты с исключенным ОНМК	50	66,8 (23- 96)	25	50		25	50	

Таблица 2 – Распределение данных о сроке выполнения КТ от начала заболевания, бассейне ОНМК, окклюзированном сегменте средней мозговой артерии у пациентов, включенных в базу данных

Группа пациентов	Количество	Срок выполнения КТ от начала заболевания			Бассейн ОНМК		Тромбоз сегмента СМА	
		Менее 6 часов	Более 6 часов	Более 24 часов	ЛСМА	ПСМА	М1	М2
Пациенты с подтвержденным ОНМК	100	69	31	71	52	48	79	21
Пациенты с исключенным ОНМК	50	50	0	28	-	-	-	-

Таблица 3 – Распределение данных о признаках ишемического инсульта, а также баллах по шкале ASPECTS у пациентов, включенных в базу данных

Группа пациентов	Количество	Признаки ишемического инсульта		ASPECTS		
		Гипертензивная артерия	РИИ	10	от 9 до 6	от 5 до 0
Пациенты с подтвержденным ОНМК	100	79	67	19	51	30
Пациенты с исключенным ОНМК	50	0	0	50	0	0

Таблица 4 – Распределение данных КТ- перфузии, о тромбэкстракции и об исходе

Группа пациентов	Количество	Данные перфузии			Проведение тромбэкстракции	Исход		
		Наличие пенумбры	Соотношение ядра инсульта			Геморрагическая трансформация инсульта	Отрицательная динамика ASPECTS в течении 24 часа	Летальный исход
			менее 40%	более 40 %				
Пациенты с ОНМК	100	100	82	18	73	33	49	20
Пациенты без ОНМК	50	0	0	0	0	0	0	0

База предназначена для обучения врачей-рентгенологов кабинетов и отделений компьютерной томографии с разным стажем и опытом в выявлении признаков ишемических изменений в острейшей и острой стадиях в бассейне средней мозговой артерии, проверки квалификации врачей-рентгенологов, тестирования систем автоматического анализа изображений, полученных методом компьютерной томографии.

На основе базы данных были сформированы 2 выборки бесконтрастных КТ-изображений головного мозга. В первую выборку вошли бесконтрастные КТ-изображения 50 пациентов с подтвержденным ОНМК и 50 пациентов, у которых данный диагноз был исключен (соотношения патология/ норма - 1:1), посредством КТ- перфузии, КТ- ангиографии брахиоцефальных артерий и сосудов головного мозга, а также проведения повторного КТ- сканирования через сутки. Были подготовлены три вариации 1 выборки для третьего (1А) и пятого этапов (1В и 1С) диссертационного исследования с измененным порядком КТ головного мозга. Во вторую выборку вошли бесконтрастные КТ - изображения 50 пациентов с подтвержденным ОНМК в б. СМА, а также окклюзией М1 сегмента

средней мозговой артерии (для дальнейшей более корректной оценки по шкале ASPECTS), по данным КТ- ангиографии брахиоцефальных артерий и сосудов головного мозга, а также КТ - перфузии (Таблица 5). Были подготовлены три вариации 2 выборки для третьего (2А) и пятого этапов (2В и 2С) диссертационного исследования с измененным порядком КТ головного мозга (Таблица 6).

Таблица 5 – Распределение пациентов в сформированных на основе базы данных выборках

Группа пациентов	Количество пациентов	Выборка 1		Выборка 2	
		50	100	50	50
Пациенты с подтвержденным ОНМК	100	50	100	50	50
Пациенты с исключенным ОНМК	50	50		0	

Таблица 6 – Этапы тестирования врачей с соотношением с использованными вариациями выборок КТ- изображений

Этап исследования			Выборка
Третий этап исследования	Оценка КТ- исследований врачами (n=21) с различным стажем и опытом в ургентной нейрорадиологии	Наличие/ отсутствие ОНМК в б. СМА, сторона поражения	1А
		Сторона поражения, конечный балл по ASPECTS	2А
Пятый этап исследования	Оценка КТ- исследований врачами (n=7) с различным опытом в оценке ишемического инсульта и стажем до 3 лет	Наличие/ отсутствие ОНМК в б. СМА, сторона поражения	1В
		Сторона поражения, зоны и конечный балл по ASPECTS	2В

Пятый этап исследования	Оценка КТ- исследований врачами (n=7) с различным опытом в оценке ишемического инсульта и стажем до 3 лет	Наличие/ отсутствие ОНМК в б. СМА, сторона поражения	1С
		Сторона поражения, зоны и конечный балл по ASPECTS	2С

Критерий включения во 1 выборку для пациентов с подтвержденным ОНМК в бассейне СМА – наличие ишемического инсульта, верифицированного с помощью КТ-ангиографии брахиоцефальных и церебральных сосудов, а также КТ-перфузии [4, 5].

Критерии исключения для данной группы пациентов:

- 1) повторное ОНМК в бассейне средней мозговой артерии;
- 2) ОНМК в бассейнах задней циркуляции и передней мозговой артерии;
- 3) тромбоз мозговых вен и венозных синусов;
- 4) хроническая окклюзия артерий передней циркуляции;
- 5) отсутствие изменения на картах КТ- перфузии.

Критерий включения в исследование в 1 выборку для пациентов, у которых диагноз ОНМК был исключен:

- 1) отсутствие КТ- признаков ишемических изменений по данным бесконтрастной КТ в бассейне средней мозговой артерии, с подтверждением отсутствия тромбоза брахицефальных и церебральных артерий и вен с помощью КТ-ангиографии, а также наличия изменений при КТ-перфузии;
- 2) отсутствие КТ- признаков ишемических изменений по данным повторного бесконтрастного КТ- исследования в бассейне средней мозговой артерии, выполненной через 24 часа.

Критерии исключения для данной группы пациентов:

- 1) наличие ОНМК и других патологических изменений, выявляемых с помощью КТ.

Критерий включения во 2 выборку – наличие ишемического инсульта в бассейне средней мозговой артерии (с тромбозом М1 сегмента), подтвержденного с помощью КТ-ангиографии брахиоцефальных и церебральных сосудов, а также карт КТ-перфузии.

Критерии исключения:

- 1) повторное ОНМК в бассейне средней мозговой артерии;
- 2) ОНМК в бассейнах задней циркуляции и передней мозговой артерии (в т.ч. острого тромбоза внутренней сонной артерии, при котором наблюдается распространение тромботических масс в переднюю мозговую артерию);
- 3) тромбоз мозговых вен и венозных синусов;
- 4) хроническая окклюзия брахицефальных артерий головного мозга передней циркуляции;
- 5) отсутствие изменения на картах КТ- перфузии.

2.2. Проблема и вопросы межэкспертного согласия врачей рентгенологов с разным стажем и опытом работы в ургентной медицине при диагностике ишемического инсульта в бассейне средней мозговой артерии

На третьем этапе диссертационного исследования было проведено два типа тестирования.

В первом приняли участие 21 врач - рентгенолог (Таблица 7), работающие либо в региональных сосудистых центрах (N=12), либо в иных медицинских учреждениях (стационарах общего профиля), не относящихся к ургентной медицине (N=9) и не имеющие постоянного опыта оценки инсульта. Врачи разделялись в группы по стажу: до 3 лет (N=7), от 3 до 8 лет (N=7) и более 8 лет (N=7).

Таблица 7 – Ранжированные данные врачей-рентгенологов первого этапа тестирования

Тип учреждения	РСЦ			СОП		
Стаж	Менее 3 лет	от 3 до 8 лет	Более 8 лет	Менее 3 лет	от 3 до 8 лет	Более 8 лет
Количество	4	4	4	3	3	3

Для оценки была задействована 1А выборка деперсонализированных данных сформированной базы, в которую были включены 50 исследований пациентов с верифицированным ишемическим инсультом, и 50 КТ- исследований пациентов, с неподтвержденным диагнозом ОНМК.

КТ- исследования пациентов с подтвержденным ОНМК в бассейне средней мозговой артерии из первой выборки были консенсуально оценены тремя врачами с экспертной квалификацией в ургентной нейрорадиологии (с опытом диагностики инсульта) и стажем более 10 лет. Врачи совместно пришли к решению, что все КТ- исследования с подтвержденным ОНМК имеют признаки ранних ишемических изменений (с ASPECTS 9 и ниже), в бассейне окклюзированной средней мозговой артерии.

При первом тестировании, в процессе анализа бесконтрастных КТ- исследований специалисты заполняли форму с ответами, интерпретируя КТ- изображения как норму или патологию, а также наличие таких признаков как сторона поражения, наличия признака гиперденсной артерии, признака нарушения дифференцировки белого и серого вещества (также включающего в себя «признак инсулярной ленты» и исчезновение границ (контрастности) базальных ганглиев), а также сглаженности корковых борозд и снижения КТ- плотности вещества головного мозга.

После предоставления бесконтрастных КТ- исследований врачам рентгенологам, просмотр КТ осуществлялся с помощью программы RadiAnt DICOM Viewer (с возможностью построения реформации изображений MPR) как

универсального инструмента, устанавливаемого на любой компьютер. Просмотр изображений осуществлялся на мониторах HP EliteDisplay S340с с разрешением 3440x1440 Мпикс, операционная система Windows 10.

Настройки окна включали как стандартные значения ширины окна в 80 HU и уровня окна в 40 HU, так и нестандартные значения с узкими шириной и уровнем окна составляющими приблизительно в 35-45 HU для ширины и 35-45 HU для уровня окна, соответственно, для оценки по шкале ASPECTS.

Эксперты вне зависимости друг от друга оценивали каждое исследование и маркировали признаки при наличии как “1”, при отсутствии как “0”. Далее все оценки были объединены независимым исследователем в общую таблицу для дальнейшей статистической обработки результатов.

Случай 1 – КТ- изображение головного мозга без введения контрастного препарата, с признаком гиперденсной артерии. Абсолютная плотность правой средней мозговой артерии составляет +63 HU (должна составлять не менее +43 HU), с соотношением плотности левой СМА - 1,2 (Рисунок 9).

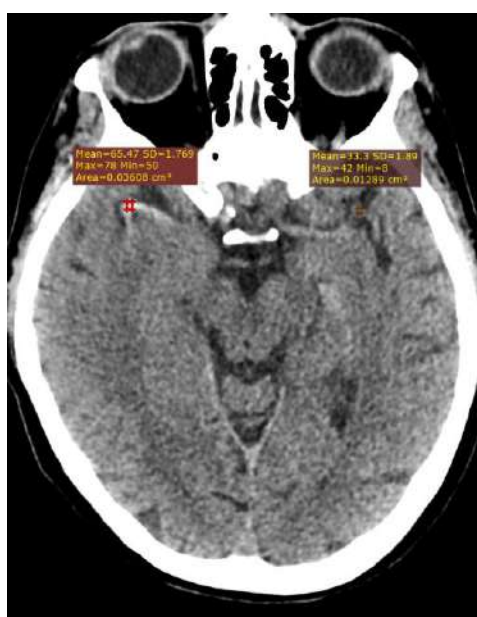


Рисунок 9 – ОНМК по ишемическому типу в бассейне правой средней мозговой артерии. Данные КТ- исследования головного мозга, случай N1.

Случай 2 – КТ- изображение головного мозга без введения контрастного препарата, с признаком исчезновения границ (контрастности) базальных ганглиев правого полушария (Рисунок 10).



Рисунок 10 – ОНМК по ишемическому типу в бассейне правой средней мозговой артерии. Данные КТ- исследования головного мозга, случай N2.

Случай 3 – КТ- изображение головного мозга без введения контрастного препарата, с признаком сглаженности корковых борозд левой лобной доли (Рисунок 11).

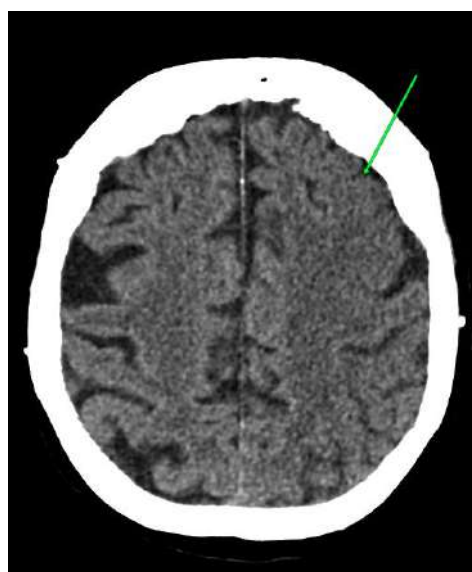


Рисунок 11 – ОНМК по ишемическому типу в бассейне левой средней мозговой артерии. Данные КТ- исследования головного мозга, случай N3.

Случай 4 – КТ- изображение головного мозга без введения контрастного препарата (Рисунок 12), с признаком снижения КТ- плотности вещества головного мозга (до +20 НУ).



Рисунок 12 – ОНМК по ишемическому типу в бассейне левой средней мозговой артерии. Данные КТ- исследования головного мозга, случай N4.

Во втором тестировании участвовали 15 врачей - рентгенологов (Таблица 8), работающих только в РСЦ, разделенные на равные группы согласно стажу: до 3 лет (N=5), от 3 до 8 лет (N=5), более 8 лет (N=5). С учетом специфики их работы и рутинности использования шкалы ASPECTS для оценки распространенности ишемических изменений в бассейне средней мозговой артерии, врачам было предложено ранжировать балл по ASPECTS по своему субъективному мнению. Для оценки была задействована 2А выборка деперсонализированных данных сформированной базы, в которую были включены 50 исследований пациентов с верифицированным ишемическим инсультом (с выявленной острой окклюзией М1 сегмента СМА).

Таблица 8 – Ранжированные данные врачей-рентгенологов второго этапа тестирования

	РСЦ		
Стаж	Менее 3 лет	от 3 до 8 лет	Более 8 лет
Количество	5	5	5

При проведении статистического анализа были использованы методы описательной статистики. Анализ выполнен с использованием программы SPSS Statistics 19 и языка программирования Python.

На основании результатов оценки первого тестирования были осуществлены расчеты таких параметров, как чувствительность, специфичность, точность для оценки наличия/отсутствия ОНМК в бассейне СМА (проведен ROC-анализ и рассчитана площадь под ROC-кривой) [4, 77].

При оценке КТ-признаков ишемического инсульта статистический анализ включал определение коэффициента межэкспертного согласия (каппа Флейсса) относительно КТ-признаков ишемического инсульта [64]. Коэффициент межэкспертного согласия (каппа Флейсса) расценивался следующим образом: случайный ($\kappa < 0$), слабый ($\kappa > 0-0,2$), незначимый ($\kappa > 0,21-0,4$), умеренный ($\kappa > 0,41-0,6$), существенный ($\kappa > 0,61-0,8$) и полное согласие ($\kappa > 0,81-1,0$) [75].

На основании результатов оценки второго тестирования статистический анализ включал определение коэффициентов межэкспертного согласия (каппа Коэна – путем попеременной оценки согласованности экспертов друг с другом и каппа Флейсса) [59]. Результаты оценки межэкспертного согласия с применением каппы Коэна (k) представлены в виде диапазона значений от минимального до максимального ($\min k - \max k$), а также среднего значения (μ). Применение двух коэффициентов для измерения межэкспертного согласия обусловлено разными вариантами статистического анализа в литературе и необходимо для корректного сравнения полученных данных с данными других исследователей. Кроме того, статистический анализ проводился с помощью дихотомического деления баллов

по шкале ASPECTS: ≤ 6 и > 6 и ≤ 7 и > 7 , поскольку данные литературы разнятся в обсуждении порогового балла ASPECTS, коррелирующего с худшим функциональным исходом, а также с риском развития внутримозгового кровоизлияния [112]. Коэффициенты межэкспертного согласия (каппа Коэна и каппа Флейсса) расценивались следующим образом: слабый ($k > 0-0,2$), незначимый ($k > 0,21-0,4$), умеренный ($k > 0,41-0,6$), существенный ($k > 0,61-0,8$) и полное согласие ($k > 0,81-1,0$) [75].

2.3. Тестирование систем автоматического анализа

На четвертом этапе диссертационного исследования для тестирования нами было выбрано три программных продукта на основе сверточных нейронных сетей, позиционирующих себя как программное устройство для компьютерной медицинской диагностики, используемое врачом для интерпретации характера и распределения патологических изменений в тканях головного мозга путем оценки по шкале ASPECT на основе данных КТ.

Критериями отбора программ для тестирования являлись:

1. Возможность тестового онлайн доступа;
2. В описании программного продукта указана функция оценки участков с ишемическим поражением тканей головного мозга в острый период (т. е. оценкой по шкале ASPECTS)
3. Заявленная точность в диагностике ишемического инсульта в бассейне средней мозговой артерии (СМА) более 75% (с учетом результатов 1 этапа исследования, описано в главе 3).

Согласно критериям включения было выбрано четыре программных продукта (2 системы отечественных разработчиков и 1 система иностранного производства).

Поскольку целью исследования ставилась общая оценка диагностических показателей имеющихся в настоящее время систем, а не оценка того или иного продукта, все программы, включенные в исследование, были анонимизированы

(А, В, С). Заявленная разработчиками точность, чувствительность и специфичность программ автоматического анализа представлены в таблице 9.

Таблица 9 – Заявленные показатели диагностической эффективности программных продуктов, участвующих в исследовании

Программный продукт	Точность	Чувствительность	Специфичность
А	97,0	99,0	94,0
В	83,0	82,0	82,0
С	80,0	90,0	70,0

Программа А позиционируется разработчиками как программное обеспечение на основе технологии искусственного интеллекта (ИИ) предназначенное для обработки КТ-изображений головного мозга при подозрении на ОНМК с целью раннего выявления ургентной патологии (ишемический инсульт) и уведомления о ней. Требования, предъявляемые к КТ - изображениям головного мозга для обеспечения оптимальной производительности программы А представлены в таблице 10.

Таблица 10 – Рекомендуемые параметры получения изображений для целей обработки с помощью алгоритмов

Область сканирования	При компьютерной томографии головы рекомендуется захватить весь головной мозг (от конвекса до уровня большого затылочного отверстия)
Размер матрицы	512x512
Толщина среза	0,625-5 мм
DFOV (поле обзора отображения по оси X/Y)	180-324 мм

В соответствии с рекомендациями ASPECTS для каждого полушария головного мозга определяется 10 областей интереса, которые обводятся на изображении желтым контуром. Области интереса, определяемые алгоритмом ASPECTS как имеющие РИИ, обводятся красным контуром и отображаются на панели отчетности ASPECTS справа от окна просмотра. Общая оценка в баллах по шкале ASPECTS отображается на панели отчетности и отражает количество областей из 10, определенных как не затронутые изменениями (показано зеленым), в то время как количество областей из 10, определенных как вовлеченные в патологический процесс, показаны красным. Кроме того, пользователь может включать или выключать области интереса наложения изображений и работать с входящим бесконтрастным КТ - изображением, запрашивая аннотацию “ASPECTS: ВКЛ/ВЫКЛ” в области просмотра (Рисунок 13).

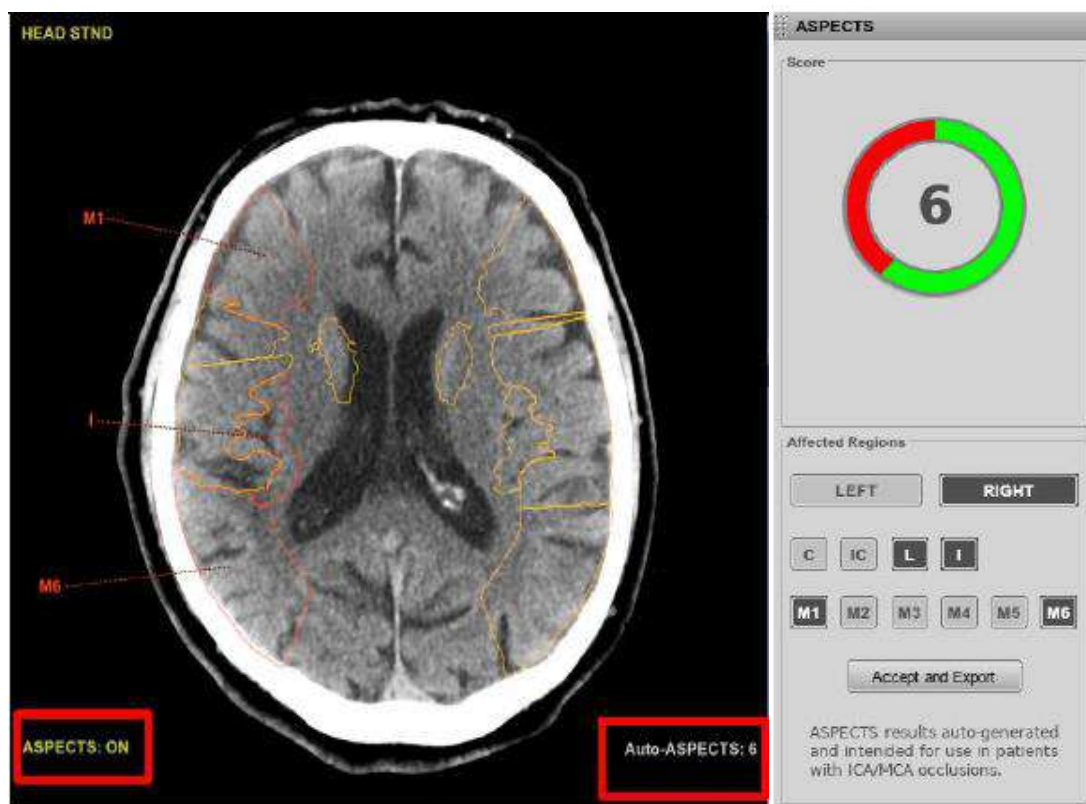


Рисунок 13 – Интерфейс и результат обработки бесконтрастного КТ - исследования головного мозга программой А

Аналогичным образом, программа А выставляет оценку 10 (т.е. области с признаками РИИ не выявлены), все области на панели отчетности отображаются в сером цвете, а все наложения областей в окне просмотра показаны желтым цветом.

Программа А была обучена на базе данных, состоящей из КТ- исследований, полученных в ходе многоцентрового клинического исследования Prove-IT (идентификатор ClinicalTrials.gov: NCT02184936) с участием пациентов в возрасте 22 лет и старше. Для аналитической валидации программы А был использован набор данных 200 пациентов, из двух клинических исследований и двух рандомизированных контролируемых исследований: Prove-IT (N = 40, идентификатор ClinicalTrials.gov: NCT02184936), INTERRSeCT (N = 59, идентификатор JamaNetwork.com: 2702146), ESCAPE (N = 16, идентификатор ClinicalTrials.gov: NCT01778335), ESCAPE-NA1 (N = 85, идентификатор ClinicalTrials.gov: NCT02930018). Тестовые данные были получены из медицинских учреждений различных географических регионов (Канада, США, ЕС, Азия), выполнены на разных компьютерных томографах (GE, Siemens, Philips, Toshiba). КТ- изображения пациентов, включенные в базу данных, предназначенных для аналитической валидации, носили репрезентативный характер с широким диапазоном степеней клинической тяжести (диапазон баллов по шкале ASPECT 0–10, средний балл по шкале ASPECTS = 8; диапазон баллов по шкале NIHSS 0–30, медианный балл по шкале NIHSS = 17) и времени с момента проявления симптомов до проведения КТ < 360 минут.

Программа В (ТУ 58.29.32-001-14161592-2022) - программное обеспечение, предназначенное для обработки и поддержки качества описания медицинских диагностических изображений пациентов с острым нарушением мозгового кровообращения. Программа определяет наличие очагов ишемии и сегментирует их, выделяя области в виде маски (красный цвет соответствует областям острой ишемии, синий - хронической). Система позволяет также разметить территории

бассейна средней мозговой артерии, определяемые на КТ при оценке по шкале ASPECTS, в виде маски (Рисунок 14).



Рисунок 14 – Пример выделения зоны интереса (области ASPECTS с признаками ранних ишемических изменений) в бассейне левой средней мозговой артерии программой В

Каждый цвет соответствует одной области:

- красный – хвостатое ядро (С),
- розовый – чечевицеобразное ядро (L),
- желтый – внутренняя капсула (IC),
- оранжевый – кора островковой доли (I),
- зеленый – передняя часть корковой области СМА (M1),
- фиолетовый – корковая область бассейна СМА, латеральное коры островковой доли (M2),
- бирюзовый – задняя часть коры СМА (M3),
- синий – передняя территория бассейна СМА, располагающаяся непосредственно выше и роstralнее M1 (M4),
- мятный – латеральная территория бассейна СМА, располагающаяся непосредственно выше и роstralнее M2 (M5),
- голубой – задняя территория бассейна СМА, располагающаяся непосредственно выше и роstralнее M3 (M6).

Программа была обучена и валидирована на наборе данных 600 КТ-исследований головного мозга мужчин и женщин старше 18 лет, сформированном в соответствии с ГОСТ Р 59921.5-2021 [18].

Программа С (регистрационное удостоверение медицинского изделия № 2013616688) предназначена для поддержки принятия врачебных решений для работы врачей – специалистов по лучевой диагностике и врачей неврологов при оценке степени поражения головного мозга при ишемическом инсульте по изображениям, полученным на рентгеновских компьютерных томографах (КТ).

Программа обеспечивает:

- регистрацию и визуализацию КТ изображений головного мозга пациента в стандарте DICOM;
- сегментирование в автоматическом режиме гиподенсивных областей на изображениях нативной КТ головы;
- разметку зон, соответствующих областям шкалы ASPECT с вычислением суммарного балла, характеризующего выявленные изменения;
- формирование заключения с оценкой объема гиподенсивной области и балла по шкале ASPECTS;
- представление пораженных областей ASPECTS в виде схематического изображения.

Рекомендуемая толщина среза для КТ- исследований > 1 мм.

Участки головного мозга, в которых предполагается наличие патологических изменений, выделяются светло-голубым цветом в виде маски. В случае их расположения в участках, соответствующих областям ASPECTS, оценивается суммарный балл по данной шкале. Дополнительный сервис размечает бежевым цветом гиподенсивные участки, с высокой вероятностью являющиеся кистозно-атрофическими изменениями (Рисунок 15).

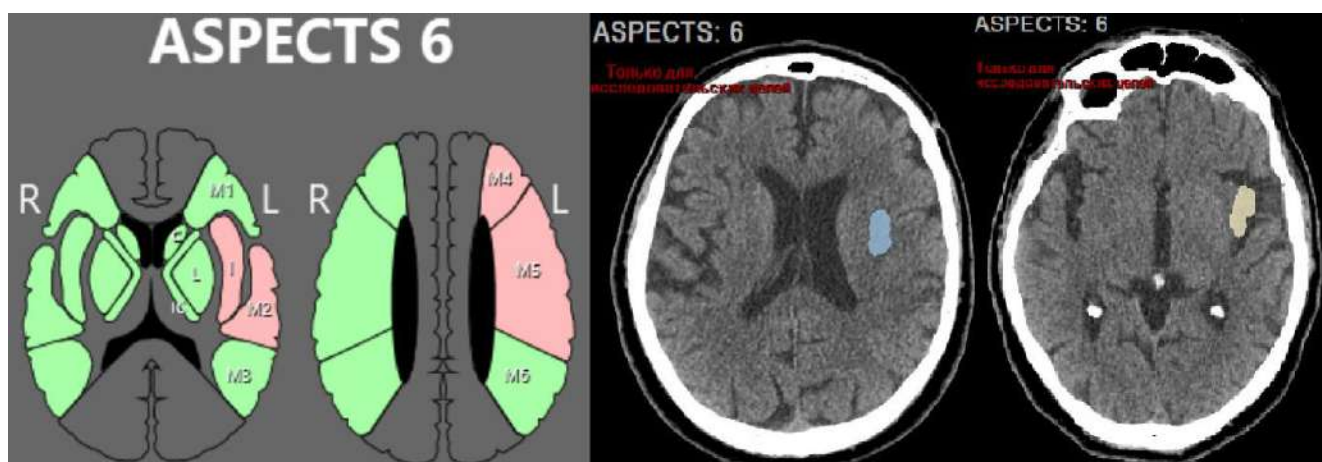


Рисунок 15 – Пример выделения зоны интереса программой С, с расчетом пораженных зон по ASPECTS в бассейне левой средней мозговой артерии, а также кистозно- атрофических изменений в левой островковой доле

Разработчиками указывается, что заключение сервиса содержит вероятностную оценку наличия ишемического инсульта и балл по шкале ASPECTS.

Ограничениями использования сервиса являются двигательные артефакты на уровне сканирования, хирургические вмешательства, проведенные на головном мозге, наличие сочетанной патологии (онкообразования), толщина срезов компьютерной томографии более 3 мм, а также технические артефакты исследования, которые не относятся к пациенту (связанные с неисправностью томографа). Стоит отметить, что программный продукт С не интерпретировал шестнадцать КТ- исследований головного мозга из 1 выборки, вследствие неустранимой технической ошибки при загрузке лучевых изображений.

Тестирование программных продуктов было произведено методом аналитической валидации – оценки эффективности системы искусственного интеллекта путем использования эталонных данных с подтверждением способности программы воспроизводимо и надежно генерировать предполагаемые технические результаты вычислений из входных данных [15, 18].

Для тестирования была использована 1 выборка изображений. Программные продукты предоставляли результаты анализа КТ- изображений в

виде балла по ASPECTS и стороны поражения. Поскольку все КТ- исследования с подтвержденным ОНМК были предварительно консенсуально оценены тремя врачами с экспертной квалификацией в ургентной нейрорадиологии (с опытом диагностики инсульта), вследствие чего они совместно пришли к решению, что все КТ- исследования с подтвержденным ОНМК имеют признаки ранних ишемических изменений (с ASPECTS 9 и ниже), то ответ программам “ASPECTS 10” трактовался как отсутствие ОНМК, ответ “ASPECTS 9 и ниже” как наличие ОНМК.

Статистический анализ выполнен с использованием программы SPSS Statistics 19 и языка программирования Python. Проводилась оценка показателей эффективности выявления ишемического инсульта в бассейне средней мозговой артерии на компьютерных томограммах (чувствительность, специфичность, отношение правдоподобия положительного результата, отношение правдоподобия отрицательного результата, количество истинно положительных, ложноположительных, ложноотрицательных и истинно отрицательных ответов, а также точность). Для сравнения этих показателей с результатами врачей, протестированных на третьем этапе диссертационного исследования, для специалистов из РСЦ со стажем менее 3 лет и более 8 лет, а также их коллег из СОП со стажем менее 3 лет, были рассчитаны дополнительные показатели диагностической эффективности (отношение правдоподобия положительного результата, отношение правдоподобия отрицательного результата, количество истинно положительных, ложноположительных, ложноотрицательных и истинно отрицательных ответов). Кроме того, были построены графики, позволяющие оценить качество бинарной классификации, – характеристические кривые (ROC-кривые) [4, 7]. При анализе все метрики оценивались в диапазоне 0–1: <0,6 — непригодно; 0,61–0,8 — требуется доработка;> 0,81 — может быть допущено к клинической валидации [15].

2.4. Комплементарная оценка компьютерно- томографических исследований системой автоматического анализа и врачами рентгенологами

На пятом этапе диссертационного исследования нами были изучены возможные варианты внедрения систем автоматического анализа компьютерных томограмм как метода диагностики ишемического инсульта в бассейне средней мозговой артерии в клиническую практику врача-рентгенолога [88].

Для клинической валидации была отобрана программа А. Программа была протестирована на 2 выборке КТ- исследований. Также были задействованы врачи со стажем работы до 3 лет с различным опытом оценки ишемического инсульта из третьего этапа диссертационного исследования (описан в главе 3). Между третьим и пятым этапом прошло одиннадцать месяцев, при предыдущем тестировании врачам не сообщались правильные ответы. Перед началом настоящего тестирования рентгенологи были ознакомлены с показателями диагностической эффективности используемой системы автоматического анализа. Распределение рентгенологов между группами 1 и 2 описаны в таблице 11.

Таблица 11 – Распределение рентгенологов по группам

	С опытом работы в РСЦ	Без опыта работы в РСЦ
1 группа	4	-
2 группа	-	3

Был проведен совместный анализ тестовых выборок, используемых для тестирования врачами рентгенологами и системой автоматического анализа с моделированием двух вариантов (временной интервал между тестированиями 2 месяца):

1. Первичный анализ бесконтрастных КТ- исследований врачом, с последующим предоставлением ему данных интерпретации системы автоматического анализа и возможной корректировкой специалиста своего ответа.

2. Первичный анализ бесконтрастных КТ- изображений системой автоматического анализа и последующей оценкой, с учетом данных интерпретации программы, врачом.

В первом варианте, специалисты в процессе анализа 1В выборки заполняли форму с ответами, интерпретируя КТ- изображения как норму или патологию, а также наличие стороны поражения. Эксперты вне зависимости друг от друга оценивали каждое исследование и маркировали их при наличии ОНМК как “1”, при отсутствии как “0”. После предоставления данных интерпретации системы автоматического анализа, они могли скорректировать свой ответ. Аналогичным образом, врачи заполняли форму при оценке КТ- исследований из выборки 2В с оценкой стороны инсульта, зон с наличием ранних ишемических изменений и конечного балла по ASPECTS. Подсчет по ASPECTS производился путем вычитания 1 балла из 10, для каждого достоверного признака ранних ишемических изменений (нарушение дифференцировки серого/белого вещества, сглаженность корковых борозд на стороне поражения, а также фокальная гиподенсивность) в каждом из регионов. Далее врачам сообщались результаты оценки алгоритма искусственного интеллекта (зоны по ASPECTS, конечный балл и сторона поражения). Врачи могли согласиться с искусственным интеллектом, не менять свое решение или изменить свой ответ с включением дополнительных результатов (зон по ASPECTS), обнаруженных программой.

Во втором варианте тестирования, проведенном через 2 месяца, при оценке КТ- исследований из выборки 1С, врачам первоначально сообщались данные интерпретации программы о наличии/ отсутствии ОНМК в бассейне средней мозговой артерии и стороны поражения. Далее врачи или соглашались с искусственным интеллектом, или предлагали свой вариант интерпретации, а также повторно заполняли форму ответов, с маркировкой исследований при наличии ОНМК как “1”, при отсутствии как “0”. Сходным образом, при оценке исследований 2С выборки врачам предоставлялись данные оценки системой стороны инсульта, зон с наличием ранних ишемических изменений и конечного

балла по ASPECTS. Рентгенолог мог как согласиться с интерпретацией программного продукта, так и составить свое окончательное мнение о стороны поражения, объема вовлечения в ишемический процесс территории кровоснабжения средней мозговой артерии по ASPECTS (Таблица 12).

Таблица 12 – Этапы тестирования врачей с соотношением с использованными вариациями выборок КТ- изображений, а также временным интервалом исследования

Этап исследования		Выборка	Временной интервал
Оценка КТ- исследований врачами (n=21) с различным стажем и опытом в ургентной нейрорадиологии	Наличие/ отсутствие ОНМК в б. СМА, сторона поражения	1А	23.05-03.06.2022
	Сторона поражения, балл по ASPECTS	2А	
Оценка КТ- исследований врачами (n=7) с различным опытом в оценке ишемического инсульта и стажем до 3 лет	Наличие/ отсутствие ОНМК в б. СМА, сторона поражения	1В	03.05-05.05.2023
	Сторона поражения, зоны и балл по ASPECTS	2В	
Оценка КТ- исследований врачами (n=7) с различным опытом в оценке ишемического инсульта и стажем до 3 лет	Наличие/ отсутствие ОНМК в б. СМА, сторона поражения	1С	01.07-04.07.2023
	Сторона поражения, зоны и балл по ASPECTS	2С	

Просмотр КТ осуществлялся с помощью программы RadiAnt DICOM Viewer (с возможностью построения реформации изображений MPR), как универсального инструмента, устанавливаемого на любой визуализационный компьютер. Просмотр изображений осуществлялся на мониторах HP EliteDisplay S340c с разрешением 3440x1440 Мпикс, операционная система Windows 10.

Статистический анализ полученных данных осуществлялся с применением IBM SPSS Statistics 19 и языка программирования Python. В связи с тем, что распределение большинства количественных параметров не подчинялось закону нормального распределения, непараметрические критерии не использовались. С учетом отсутствия достоверной разницы между результатами коэффициентов согласия каппой Коэна и каппой Флейса на третьем этапе исследования, оценка проводилась только с определением каппы Флейса. Статистический анализ при определении согласованности в отношении оценки ишемического инсульта по ASPECTS включал определение коэффициента межэкспертного согласия (каппа Флейсса) [64]. Коэффициент межэкспертного согласия расценивался следующим образом: случайный ($\kappa < 0$), слабый ($\kappa > 0-0,2$), незначимый ($\kappa > 0,21-0,4$), умеренный ($\kappa > 0,41-0,6$), существенный ($\kappa > 0,61-0,8$) и полное согласие ($\kappa > 0,81-1,0$) [75].

ГЛАВА 3. РЕЗУЛЬТАТЫ КТ- ДИАГНОСТИКИ ИШЕМИЧЕСКОГО ИНСУЛЬТА В БАССЕЙНЕ СРЕДНЕЙ МОЗГОВОЙ АРТЕРИИ ВРАЧАМИ РЕНТГЕНОЛОГАМИ В ЗАВИСИМОСТИ ОТ СТАЖА И ОПЫТА РАБОТЫ В УРГЕНТНОЙ МЕДИЦИНЕ

В ходе изучения диагностической эффективности врачей-рентгенологов и их межэкспертной согласованности в отношении оценки по ASPECTS на третьем этапе диссертационного исследования было проведено два тестирования.

В первом тестировании приняли участие двадцать один эксперт. Врачи, первоначально, были распределены на две группы: специалисты ургентной медицины (РСЦ), и их коллеги из стационаров общего профиля (СОП), имеющие малый опыт в диагностике ишемического инсульта в острейшей и острой стадиях. Внутри группы эксперты делились в зависимости от стажа работы: более 8 лет, от 3 до 8 лет, и до 3 лет - соотношение врачей РСЦ и СОП равнялось 4/3 в каждой группе, разделенной по стажу работы [4].

Экспертам предлагалось оценить наличие/отсутствие ОНМК по ишемическому типу на основе изображений бесконтрастной КТ головного мозга. Также врачи должны были установить наличие таких признаков как снижение рентгеновской плотности вещества головного мозга, сглаженность корковых борозд, нарушение дифференцировки серого/белого вещества, а также гиперденсивность средней мозговой артерии и ее дистальных отделов (признак или «симптом точки») [4].

Результаты первого тестирования с оценкой диагностических показателей врачей с различным стажем и опытом в ургентной нейрорадиологии представлены в таблице 13.

Таблица 13 – Результативность выявления ишемических изменений в зависимости от стажа и наличия опыта работы в ургентной нейрорадиологии

Стаж	Специализация	Чувствительность, %	Специфичность, %	Точность, %
Более 8 лет	РСЦ	90,5	97,0	93,8
	СОП	85,3	79,3	82,3
3-8 лет	РСЦ	92,5	95,0	91,3
	СОП	83,3	86,7	85,0
До 3 лет	РСЦ	84,0	92,5	88,3
	СОП	57,3	92,7	75,0

По данным нашего исследования, существует прямая зависимость между стажем работы врачей РСЦ и такими показателями эффективности выявления ОНМК, как точность и специфичность. Для врачей из СОП данная корреляция не сохранялась. Стоит отметить, что определяется скачок количества правильных интерпретаций КТ-изображений у экспертов каждой группы со стажем более 3 лет. Исходя из данных статистических результатов, мы можем сделать вывод, что диагностическая эффективность врачей зависит как от стажа, так и от опыта работы в ургентной нейрорадиологии [4].

Уровень гипер- и гиподиагностики у врачей РСЦ со стажем от 3 лет ниже, чем у специалистов РСЦ со стажем менее 3 лет и врачей СОП в целом. Стоит отметить, что все врачи, участвующие в тестировании были более склонны к гипердиагностике, эта тенденция более выражена среди специалистов СОП со стажем менее 3 лет. Также, исходя из вышеописанных результатов, стаж специалистов, не имеющих опыта работы в ургентной медицине, не влияет на их диагностическую эффективность [4]. Для оценки взаимосвязи чувствительности и

специфичности были построены ROC-кривые с вычислением площади под кривой (AUC) для всех анализируемых групп (Рисунок 16).

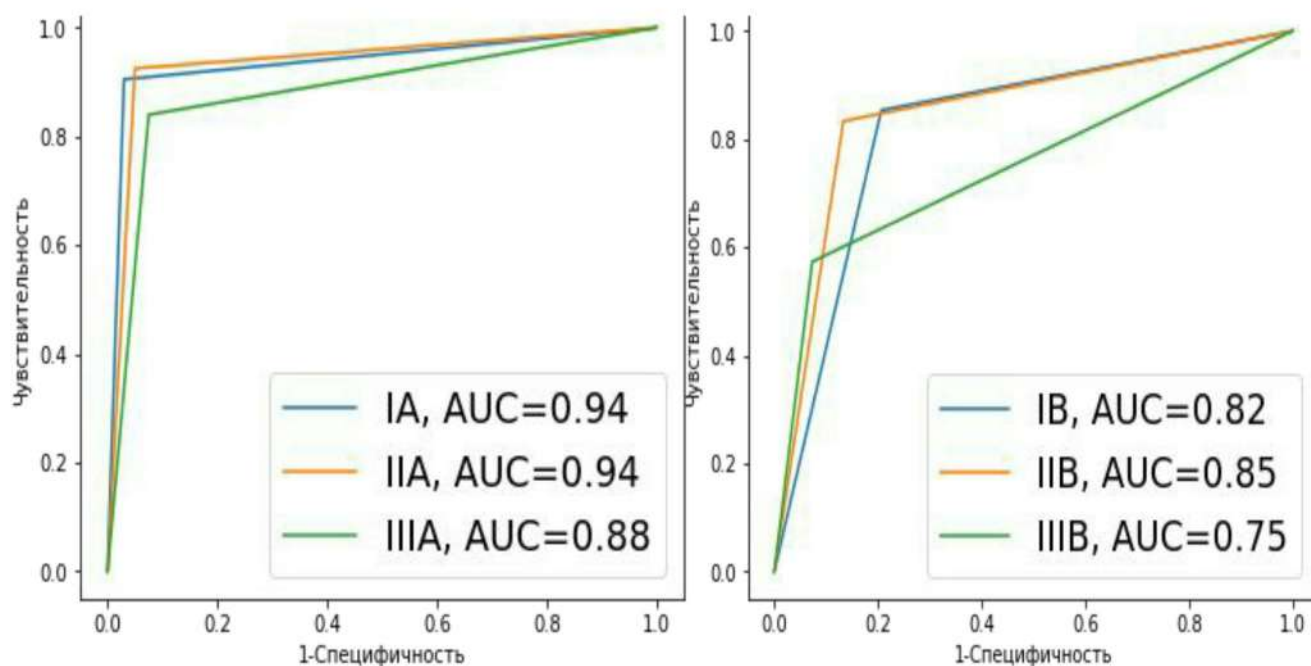


Рисунок 16 – Кривые ошибок (ROC-кривые) и расчет площадей под ними для оценки диагностической эффективности врачей - рентгенологов регионального сосудистого центра (А), и стационаров общего профиля (В) с разным стажем работы: I - более 8 лет, II - 3-8 лет, III - до 3 лет

Для оценки согласованности между исследуемыми группами специалистов в отношении КТ-признаков ишемического инсульта в бассейне СМА была рассчитана каппа Флейсса (Таблица 14).

Таблица 14 – Согласованность рентгенологов с разным стажем и опытом в ургентной нейрорадиологии в отношении КТ-признаков ишемического инсульта в бассейне средней мозговой артерии

Стаж		Более 8 лет		3-8 лет		До 3 лет	
		РСЦ (I)	СОП (I)	РСЦ (II)	СОП (II)	РСЦ (III)	СОП (III)
Специализация		РСЦ (I)	СОП (I)	РСЦ (II)	СОП (II)	РСЦ (III)	СОП (III)
Капша Флейсса	Гиперденсная средняя мозговая артерия	0,750	0,406	0,697	0,631	0,591	0,095
		РСЦ	0,684		СОП	0,318	
	Нарушение дифференцировки серого и белого вещества	0,671	0,519	0,626	0,530	0,599	0,114
		РСЦ	0,629		СОП	0,378	
	Сглаженность корковых борозд	0,689	0,254	0,416	0,454	0,433	-0,020
		РСЦ	0,525		СОП	0,244	
	Снижение КТ-плотности вещества головного мозга	0,535	0,495	0,493	0,530	0,479	0,077
		РСЦ	0,529		СОП	0,449	

I – врачи со стажем более 8 лет; II – врачи с опытом работы от 3 до 8 лет;

III – врачи с опытом работы до 3 лет

Общее согласие специалистов РСЦ по всем признакам колебалось от умеренного до существенного, согласие между врачами СОП ниже - от слабого до умеренного.

Межэкспертное согласие относительно признака гиперденсной средней мозговой артерии было существенным среди врачей РСЦ с опытом работы более 3 лет (с более высокими результатами при опыте выше 8 лет) и умеренным среди экспертов со стажем до 3 лет. Среди врачей СОП межэкспертная согласованность была наибольшей среди врачей со стажем от 3 до 8 лет и являлась существенной. Среди врачей со стажем более 8 лет согласие было умеренным, среди врачей со стажем до 3 лет слабым [4].

Согласованность в группах, разделенных по опыту, была существенной среди врачей со стажем более 3 лет как из РСЦ, так и из СОП, и незначимой среди экспертов со стажем менее 3 лет.

Межэкспертная согласованность относительно признака нарушения дифференцировки серого и белого вещества была аналогичной признаку гиперденсивности СМА, за исключением того, что врачи без опыта работы в ургентной медицине и стажем от 3 до 8 лет показали умеренное согласие относительно данного признака (в сравнении с существенным уровнем касательно признака гиперденсивности СМА) [4].

В отношении признака снижения КТ- плотности вещества головного мозга согласие среди всех специалистов, за исключением врачей без опыта в экстренной нейрорадиологии и стажем менее 3 лет, было умеренным. У специалистов СОП со стажем работы менее 3 лет согласованность была слабая. Стоит обратить внимание на существенное согласие среди всех специалистов РСЦ по данному признаку и умеренное среди врачей СОП, что говорит о влиянии только релевантного опыта на правильную трактовку специфических признаков [4].

Сглаженность корковых борозд являлась самым неоднозначным признаком при оценке согласия между специалистами. Существенная согласованность отмечалась только среди специалистов РСЦ со стажем более 8 лет. Умеренное

согласие регистрировалось между врачами со стажем от 3 до 8 лет в обеих группах и среди специалистов РСЦ со стажем менее 3 лет. Незначимая согласованность была выявлена среди врачей СОП со стажем более 8 лет, согласие среди их коллег со стажем менее 3 лет было случайным [4].

По итогам первого тестирования можно сделать вывод, что врачи РСЦ обладают более высокими показателями диагностической эффективности в отношении выявления ОНМК в бассейне средней мозговой артерии и преобладающей согласованностью в отношении оценки ранних ишемических изменений, по сравнению с коллегами из СОП.

Дальнейший ход исследования был направлен на определение межэкспертной вариативности между врачами непосредственно специализирующихся на ургентной нейрорадиологии. При втором тестировании 50 анонимизированных исследований были представлены 15 рентгенологам с опытом работы от 1 до 10 лет в стационаре скорой медицинской помощи. В исследовании приняли участие пять врачей со стажем более 8 лет (I группа), пять врачей со стажем работы от 3 до 8 лет (II группа) и пять врачей со стажем до 3 лет (III группа). Врачи вне зависимости друг от друга оценивали каждое исследование и определяли сторону поражения, а также балл по ASPECTS. Результаты подвергались дихотомическому делению (≤ 7 и > 7 и ≤ 6 и > 6), для дальнейшего статистического анализа (Таблица 15).

Таблица 15 – Статистический анализ межэкспертной согласованности (к Коэна и к Флейсса) при оценке по ASPECTS врачами – рентгенологами РСЦ при дихотомическом делении (≤ 6 и > 6)

ASPECTS 6				Слабое	Незначимое	Умеренное	Существенное	Полное согласие
Стаж	Каппа Флейсса	min k – max k (каппа Коэна)	μ (каппа Коэна)	0–0,2	0,21–0,4	0,41–0,6	0,61–0,8	0,81–1
Более 8 лет	0,366	0,125–0,633	0,36	20 %	40 %	30 %	10 %	0
От 3 до 8 лет	0,452	0,31–0,625	0,46	0	30 %	60 %	10 %	0
Менее 3 лет	0,462	0,31–0,646	0,46	0	30 %	60 %	10 %	0
Общий коэффициент межэкспертного согласия (каппа Коэна)								
	0,391	0,011–0,789	0,389	15,2 %	34 %	40 %	10,5 %	0

При дихотомическом делении (≤ 6 и > 6) коэффициент межэкспертного согласия (каппа Коэна, k) в I группе колебался между 0,125–0,633 (был более разнороден, чем во II и III группах) и в большей степени (40 %) был незначимым, а также включил в себя процент (20 %) результатов каппы Коэна, соответствующей слабому согласию (во II и III группах таких низких результатов не было зарегистрировано). Во II группе каппа Коэна колебалась между 0,31–0,625, составила в среднем 0,46 и в большей степени соответствовала умеренному согласию. Результаты III группы (0,31–0,646) были практически однозначны с результатами II группы [5]. Также каппа Коэна между специалистами

варьировалась от 0,011 (слабое согласие) до 0,789 (существенное согласие). Среди всех специалистов преимущественно согласие было или умеренным, или незначимым [5].

Коэффициент межэкспертного согласия (каппа Коэна, k) был выше среди во II и III группах и в среднем соответствовал значению, равному 0,48, что является умеренным показателем. Худшее согласие наблюдалось среди I и III группами, каппа Коэна составила 0,26, что является незначимым согласием (также включая большой процент слабой согласованности между специалистами) (Таблица 16) [5].

Таблица 16 – Межэкспертное согласие (каппа Коэна) среди групп специалистов при дихотомическом делении (≤ 6 и > 6)

ASPECTS 6			Слабый	Незначимый	Умеренный	Существенный	Полное согласие
Врачи	min k – max k	μ	0–0,2	0,21–0,4	0,41–0,6	0,61–0,8	0,81–1
I и II группы							
1	0,107–0,789	0,47	20 %	20 %	20 %	40 %	0
2	0,12–0,559	0,34	20 %	40 %	40 %	0	0
3	0,312–0,545	0,41	0	40 %	60 %	0	0
4	0,011–0,286	0,13	80 %	20 %	0	0	0
5	0,336–0,719	0,47	0	40 %	40 %	20 %	0
		0,37					

II и III группы							
6	0,351–0,545	0,46	0	20 %	80 %	0	0
7	0,429–0,694	0,54	0	0	60 %	40 %	0
8	0,483–0,634	0,52	0	0	80 %	20 %	0
9	0,351–0,694	0,47	0	60 %	20 %	20 %	0
10	0,247–0,582	0,4	0	40 %	60 %	0	0
		0,48					
I и III группы							
11	0,043–0,428	0,2	40 %	40 %	20 %	0	0
12	0,093–0,321	0,2	40 %	60 %	0	0	0
13	0,096–0,307	0,19	40 %	60 %	0	0	0
14	0,163–0,662	0,4	20 %	20 %	40 %	20 %	0
15	0,118–0,513	0,32	20 %	60 %	20 %	0	0
		0,26					

I – врачи со стажем более 8 лет; II – врачи с опытом работы от 3 до 8 лет;

III – врачи с опытом работы до 3 лет.

Каппа Флейсса при дихотомическом делении (≤ 6 и >6) среди трех групп составляла 0,391. В I группе каппа Флейсса была незначимой (0,366), во II и III группах умеренной (0,452 и 0,462 соответственно) [5].

Следует отметить, что были выявлены минимальные различия показателей межэкспертного согласия (каппа Коэна и каппа Флейсса), и их результаты находились в одном диапазоне, демонстрируя незначимый уровень согласия между собой при оценке по шкале ASPECTS врачами со стажем более 8 лет в ургентной нейрорадиологии и умеренный уровень согласия между врачами со стажем работы до 8 лет [5].

При дихотомическом делении (≤ 7 и > 7) коэффициент межэкспертного согласия (каппа Коэна) в I группе колебался между 0,294–0,588 и в большей степени (60 %) был умеренным, а также включил в себя процент (40 %) результатов каппы Коэна, соответствующей незначимому согласию (что выше, чем результаты при дихотомическом делении (≤ 6 и > 6)). Во II группе каппа Коэна варьировалась между 0,27–0,518 и в большей степени соответствовала незначимому согласию. Результат III группы (0,271–0,64) был наиболее высоким, составил 0,469 и соответствовал умеренному согласию (Таблица 17) [5].

Таблица 17 – Статистический анализ межэкспертной согласованности (к Коэна и к Флейсса) при интерпретации и использовании ASPECTS врачами – рентгенологами РСЦ при дихотомическом делении (≤ 7 и > 7)

ASPECTS 7				Слабый	Незначимый	Умеренный	Существенный	Полное согласие
Стаж	Каппа Флейсса	min k – max k (каппа Коэна)	μ (каппа Коэна)	0–0,2	0,21–0,4	0,41–0,6	0,61–0,8	0,81–1
Более 8 лет	0,439	0,294–0,588	0,446	0	40 %	60 %	0	0
От 3 до 8 лет	0,384	0,27–0,518	0,392	0	50 %	50 %	0	0
Менее 3 лет	0,466	0,271–0,64	0,469	0	30 %	50 %	20 %	0
Общий коэффициент межэкспертного согласия								
	0,376	0,007–0,8	0,39	12 %	40 %	46 %	2 %	0

Между специалистами каппа Коэна колебалась от 0,007 (слабое согласие) до 0,8 (существенное). Среди всех специалистов согласие было либо умеренным, или незначимым практически в равной степени [5].

Коэффициент межэкспертного согласия (каппа Коэна) был выше среди II и III групп и соответствовал значению, равному 0,45, что указывает на умеренную согласованность. Самые низкие показатели согласованности наблюдались среди I и III группами, каппа Коэна составила 0,31, что является незначимым согласием (данный результат не значительно отличался от k Коэна между I и II группами – 0,34) (Таблица 18) [5].

Таблица 18 – Межэкспертное согласие (каппа Коэна) среди групп при дихотомическом делении (≤ 7 и > 7)

ASPECTS 7			Слабый	Незначимый	Умеренный	Существенный	Полное согласие
Врачи	min k – max k (каппа Коэна)	μ	0–0,2	0,21–0,4	0,41–0,6	0,61–0,8	0,81–1
I и II группы							
1	0,331–0,504	0,42	0	40 %	60 %	0	0
2	0,154–0,485	0,32	20 %	60 %	20 %	0	0
3	0,331–0,541	0,4	0	60 %	40 %	0	0
4	0,128–0,412	0,2	80 %	0	20 %	0	0
5	0,326–0,469	0,38	0	80 %	20 %	0	0
		0,34					

II и III группы							
6	0,451–0,6	0,68	0	0	100 %	0	0
7	0,217–0,8	0,43	0	20 %	80 %	0	0
8	0,313–0,595	0,48	0	20 %	80 %	0	0
9	0,131–0,595	0,32	20 %	40 %	40 %	0	0
10	0,124–0,504	0,37	0	40 %	60 %	0	0
		0,45					
I и III группы							
11	0,007–0,244	0,14	80 %	20 %	0	0	0
12	0,118–0,388	0,29	20 %	80 %	0	0	0
13	0,099–0,374	0,27	20 %	80 %	0	0	0
14	0,231–0,587	0,41	0	40 %	60 %	0	0
15	0,2–0,6	0,44	20 %	0	80 %	0	0
		0,31					

I – врачи со стажем более 8 лет; II – врачи с опытом работы от 3 до 8 лет; III – врачи с опытом работы до 3 лет.

Каппа Флейсса при дихотомическом делении (≤ 7 и > 7) среди трех групп составляла 0,376. В I и III группах каппа Флейсса была умеренной (0,439 и 0,466 соответственно), а во II группе незначимой (0,384). Существенных различий численных показателей межэкспертного согласия (k Козна и k Флейсса) при дихотомическом делении (≤ 7 и > 7) также выявлено не было, поэтому в

дальнейшем мы производили анализ межэкспертной согласованности только с помощью каппы Флейсса [5].

Коэффициенты межэкспертного согласия каппа Коэна и каппа Флейсса были выше при дихотомическом делении (≤ 6 и > 6) [4].

При дихотомическом делении (≤ 6 и > 6) коэффициент межэкспертного согласия (каппа Коэна) преобладал во II и III группах (0,46), при дихотомическом делении (≤ 7 и > 7) был выше у III группы экспертов, что указывает о тенденции среди специалистов со стажем менее 3 лет завышать балл по ASPECTS [5].

Коэффициент согласия (k Коэна) был выше при сравнении результатов II и III групп (как при дихотомическом делении [≤ 6 и > 6], так и при [≤ 7 и > 7]), в то время как согласованность между I и III группами была наименьшей при обоих дихотомических делениях [5].

Данные результаты показывают, что согласованность экспертных оценок врачей ургентной нейрорадиологии не зависит от стажа специалиста, и в условиях данного исследования рентгенологи с большим опытом показали достаточно разнородные результаты, но однозначно менее высокие результаты, чем эксперты с наименьшим опытом [5].

Таким образом, выявлено влияние стажа работы в РСЦ врачей ургентной нейрорадиологии как на диагностическую эффективность при выявлении ОНМК в бассейне средней мозговой артерии, так и на согласованность в отношении большинства КТ-признаков ишемического инсульта. Однако стаж специалистов с экспертной квалификацией не коррелирует с уровнем межэкспертной согласованности в отношении оценки по ASPECTS, что отражает не высокую воспроизводимость данной шкалы.

ГЛАВА 4. РЕЗУЛЬТАТЫ ТЕСТИРОВАНИЯ СИСТЕМ АВТОМАТИЧЕСКОГО АНАЛИЗА И ИХ ВЛИЯНИЕ НА ПРОЦЕСС ПРИНЯТИЯ ВРАЧЕБНОГО РЕШЕНИЯ

На четвертом этапе диссертационного исследования для тестирования было выбрано три программы автоматического анализа ишемического инсульта в бассейне средней мозговой артерии, созданных на основе методов искусственного интеллекта, определяющих себя как комплексы поддержки принятия решений при оценке КТ изображений.

Критериями отбора программ для тестирования являлись:

1. Возможность тестового онлайн-доступа.
2. В описании программного продукта указана функция выявления ранних ишемических изменений и их оценка по шкале ASPECTS.
3. Заявленная точность в диагностике ишемического инсульта в бассейне средней мозговой артерии (СМА) более 75% (по данным третьего этапа диссертационного исследования, как минимальной точности среди тестируемых врачей).

Всего, согласно критериям включения, было выбрано три программных продукта (две системы отечественных разработчиков и одна система иностранного производства).

Поскольку целью исследования ставилась общая сравнительная оценка диагностических показателей, имеющихся в настоящее время систем автоматического анализа, а не оценка того или иного продукта, все программы, включенные в исследование, были анонимизированы (А, В, С).

Методом аналитической валидации было произведено тестирование программных обеспечений на эталонном наборе данных (выборка 1), подготовленном в соответствии с научной задачей, а также зарегистрированном в соответствии с регламентом [7, 15]. Набор данных КТ-изображений был

аналогичен выборке изображений третьего этапа исследования, в котором оценивалась диагностическая эффективность врачей рентгенологов с различным стажем и опытом в выявлении ишемического инсульта в бассейне средней мозговой артерии.

При статистическом анализе все метрики оценивались в диапазоне 0–1: <0,6 — непригодно; 0,61–0,8 — требуется доработка;> 0,81 — может быть допущено к клинической валидации [15]. Результаты анализа данных 1 выборки представлены в таблице 19.

Таблица 19 – Сравнительная характеристика систем автоматического анализа при оценке КТ-изображений

	Программа А	Программа В	Программа С
a (TP) истинно положительные	44	30	22
b (FP) ложноположительные	27	5	9
c (FN) ложноотрицательные	6	20	18
d (TN) истинно отрицательные	23	45	35
Не было определено	0	0	16
Чувствительность	0,88	0,60	0,55
Специфичность	0,46	0,90	0,80
Отношение правдоподобия положительного результата	1,63	6,00	2,69
Отношение правдоподобия отрицательного результата	0,26	0,44	0,57
Точность	0,67	0,75	0,68

В исследовании все отобранные системы автоматического анализа имели точность в диапазоне от 0,67 до 0,75, что, согласно клиническим рекомендациям [15], свидетельствует о необходимости их дальнейшей доработки. Наибольшую точность (параметр, определяющий количество правильно определенных суждений) продемонстрировала программа В (0,75). Специфичность программы В (0,90) также была выше по сравнению с результатами остальных программ (0,46; 0,80) [2].

Программа А продемонстрировала высокий показатель по определению истинно положительных результатов (44), при этом специфичность этой программы достаточно низкая (0,46), программа определила 27 ложноположительных результатов, что является признаком гипердиагностики и переобучения. Точность данной программы ниже точности результатов других программ. При этом программа продемонстрировала высокую чувствительность, что влияет на снижение риска гиподиагностики [2].

Программа В имела самые высокие показатели специфичности (0,90) при невысокой чувствительности (0,60), что позволит, при ее доработке и достижении необходимых суммарных параметров диагностической точности, рекомендовать данный алгоритм как метод первичной оценки с быстрой сортировкой пациентов с предполагаемым инсультом [2].

Ограничением оценки диагностической эффективности одного из алгоритмов (С) стало невозможность интерпретации им 16 КТ- исследований головного мозга из представленного независимого набора данных за счет неустранимой технической ошибки при загрузке лучевых изображений. В свою очередь, точность программы С была незначительно выше результатов программы А, она продемонстрировала большое количество ложноотрицательных результатов, что свидетельствует о ее склонности к гиподиагностике [2].

Низкая чувствительность программ является параметром, показывающим низкую эффективность при определении пациентов с ОНМК, так как часть результатов будет классифицироваться как отрицательные. Таким образом,

алгоритмам В и С требуется дальнейшая доработка (с дальнейшим обучением) в целях повышения эффективности анализа [2].

Отношение правдоподобия положительного результата показывает, во сколько раз выше вероятность получить положительный результат пациентов с ОНМК, чем у здоровых пациентов. Наивысшего результата достигла программа В, что подтверждают высокие показатели специфичности. В отношении правдоподобия отрицательного результата самые высокие показатели продемонстрировала программа С, что является результатом низкой чувствительности программы [2].

Для оценки взаимосвязи чувствительности и специфичности были построены ROC-кривые с вычислением площади под кривой (AUC) для всех анализируемых групп (Рисунок 17). Анализ показателей выявил, что только результаты, полученные при помощи программы В дают значения ($AUC = 0,75$), близкие к рекомендованным параметрам, согласно клиническим рекомендациям, предсказания других программ ниже [2].

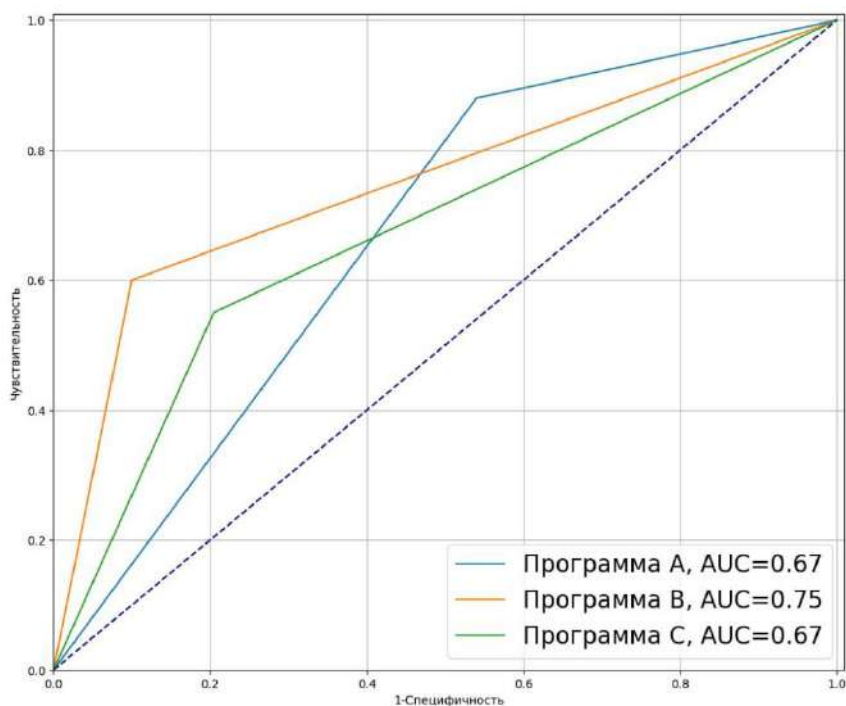


Рисунок 17 – Кривые ошибок (ROC-кривые) и расчет площадей под ними для оценки диагностической эффективности всех анализируемых программ

По результатам исследования ни одна из рассмотренных систем автоматического выявления инсульта на основе КТ-изображений не достигла пороговых значений точности, необходимых для дальнейшей клинической валидации, что свидетельствует о необходимости их дальнейшей доработки [2].

Точность всех систем автоматического анализа была значительно ниже результатов врачей с экспертной квалификацией (0,94), и также ниже результатов оценки врачей со стажем менее 3 лет (0,88) из РСЦ, протестированных на третьем этапе диссертационного исследования (Таблица 20).

Таблица 20 – Сравнительная характеристика систем автоматического анализа и рентгенологов при оценке КТ-изображений

	A	B	C	D*	F*	E*
a (TP) истинно положительные	44	30	22	45	42	29
b (FP) ложноположительные	27	5	9	2	4	4
c (FN) ложноотрицательные	6	20	18	5	8	21
d (TN) истинно отрицательные	23	45	35	48	46	46
Не было определено	0	0	16	0	0	0
Чувствительность	0,88	0,60	0,55	0,90	0,84	0,57
Чувствительность, %	88,00	60,00	55,00	90,50	84,00	57,30
Специфичность	0,46	0,90	0,80	0,97	0,93	0,93
Специфичность, %	46,00	90,00	80,00	97,00	92,50	92,70
Отношение правдоподобия положительного результата	1,63	6,00	2,69	30,17	11,20	7,82
Отношение правдоподобия отрицательного результата	0,26	0,44	0,57	0,10	0,17	0,46

Точность	0,67	0,75	0,68	0,94	0,88	0,75
Точность, %	67,00	75,00	68,00	93,80	88,30	75,00

* D - врачи РСЦ с опытом более 8 лет, F - врачи РСЦ с опытом менее 3 лет, E - врачи СОП с опытом менее 3 лет

Один из алгоритмов (B) имел сравнительную точность с врачами со стажем менее 3 лет и без опыта работы в ургентной нейрорадиологии (0,75). Наибольшее значение отношения правдоподобия положительного результата было выявлено также у врачей с экспертной квалификацией (со стажем от 8 лет в РСЦ), при этом данный показатель в 5 раз превышал наилучший результат у программ (6,00 против 30,17). Значения правдоподобия положительного результата у групп врачей со стажем менее 3 лет вне зависимости от опыта в РСЦ также выше (11,20 с опытом, 7,82 без опыта), чем у программ (A-1,63, B-6,00, C- 2,69) [2].

Для оценки взаимосвязи чувствительности и специфичности были построены ROC-кривые с вычислением площади под кривой (AUC) для всех анализируемых групп. Анализ показателей выявил, что только результаты, полученные при помощи программы B дают значения (AUC = 0,75), близкие к показателям врачей без опыта в ургентной медицине (AUC = 0,75), точность предсказаний других программ ниже (Рисунок 18).

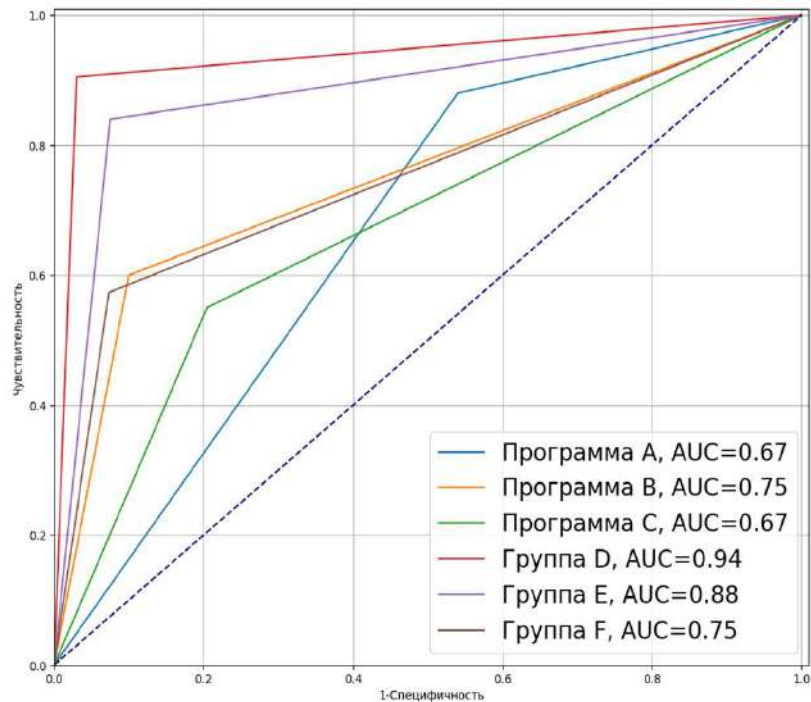


Рисунок 18 – Кривые ошибок (ROC-кривые) и расчет площадей под ними для оценки диагностической эффективности всех анализируемых программ, врачей из РСЦ со стажем более 8 лет, и врачей со стажем до 3 лет из РСЦ и СОП

Однако следует отметить, что несмотря на высокий показатель ложноположительных ответов, программа А практически не уступает результатам врачей с экспертной квалификацией относительно истинно положительных и ложноотрицательных ответов, что позволяет судить о ее низком уровне гиподиагностики. У врачей со стажем менее 3 лет и без опыта работы в РСЦ по сравнению с программой А ложноотрицательных ответов было больше в 3,5 раза, а истинно положительных ответов меньше в 1,5 раза. Относительно этих показателей врачи со стажем до 3 лет из РСЦ также несколько уступали программе А. Это говорит, что потенциально программа А может иметь положительный эффект для снижения пропуска патологии у молодых специалистов со стажем до 3 лет вне зависимости от опыта в оценке инсульта [2].

Таким образом, в настоящее время результаты систем автоматического выявления ишемических изменений на КТ- изображениях не сопоставимы с

усредненными данными результатов врачей-рентгенологов с опытом оценки ОНМК по показателям диагностической эффективности.

Большинство программных продуктов показывают высокие показатели специфичности и невысокие показатели чувствительности, что свидетельствует о нечастых случаях гипердиагностики и большого количества случаев гиподиагностики. Для более достоверного понимания диагностических возможностей данных программных продуктов следует продолжать клинические испытания как методом аналитической валидации на различных выборках, так и методом клинической валидации.

ГЛАВА 5. ВЛИЯНИЕ ВНЕДРЕНИЯ СИСТЕМ АВТОМАТИЧЕСКОГО АНАЛИЗА НА ПРОЦЕСС ПРИНЯТИЯ ВРАЧЕБНОГО РЕШЕНИЯ

На пятом этапе исследования были изучены возможные варианты внедрения систем автоматического анализа компьютерно- томографических изображений как метода поддержки принятия врачебных решений при диагностике ишемического инсульта в бассейне средней мозговой артерии.

Согласно Obuchowski N. A. и соавт. [88] существует четыре варианта использования алгоритмов ИИ в медицинской визуализации: (1) модель первого или параллельного чтения, когда алгоритм ИИ сначала оценивает изображение, а затем результаты этой оценки просматриваются человеком, (2) модель второго чтения, когда алгоритм ИИ применяется после того, как человек интерпретировал данные для возможности включения дополнительных результатов, обнаруженных ИИ, (3) модель сортировки, когда алгоритм распределяет случаи в соответствии с наличием (в процентном отношении) подозреваемой патологии (4), модель предварительного скрининга, когда алгоритм ИИ применяется к набору изображений для выявления нормы с генерированием клинического отчета и дальнейшей оценкой оставшихся (классифицируемых как “не норма”) случаев врачом.

С целью изучения внедрения первых двух вариантов использования систем автоматического анализа мы провели два варианта совместного тестирования врачей со стажем до 3 лет и алгоритма искусственного интеллекта, отобранного на четвертом этапе диссертационного исследования. Предметом изучения стала программа А, поскольку, несмотря на недостаточный уровень точности, данная система показала результаты позволяющие судить о низком уровне гиподиагностики ишемических изменений (чувствительность 88 %, истинно положительные ответы - 44 из 100, ложноположительные ответы 27 из 100). Анализ бесконтрастных КТ-изображений данным программным продуктом

выполнялся с функцией автоматической оценки участков с ишемическим поражением вещества головного мозга в острый период (т. е. оценкой по шкале ASPECTS). Также в исследование были включены семь рентгенологов со стажем работы менее 3 лет. Врачи были разделены на две группы согласно опыту в КТ-диагностики ишемического инсульта: специалисты из РСЦ (N=4) и их коллеги из СОП (N=3).

В первом варианте тестирования, в процессе исследования внедрения модели второго чтения использовались выборки 1В и 2В. На КТ-изображениях выборки 1В врачам предлагалось оценить наличие/отсутствие ОНМК по ишемическому типу, далее врачам сообщались результаты оценки алгоритма искусственного интеллекта: наличие (ASPECTS 9 и ниже)/отсутствие (ASPECTS 10) инсульта, а также сторона поражения. Врачи могли как согласиться с искусственным интеллектом, так и не менять свое решение. На КТ-изображениях выборки 2В специалисты определяли сторону инсульта и распространенность ишемических изменений по ASPECTS. Далее врачам сообщались результаты оценки алгоритма искусственного интеллекта: балл по ASPECTS с указанием пораженных зон, а также сторона поражения. Врачи могли согласиться с искусственным интеллектом, не менять свое решение или изменить свое решение с включением дополнительных результатов, обнаруженных программой.

Результаты первого этапа тестирования представлены в таблицах 21 и 22.

Таблица 21 – Показатели диагностической эффективности врачей с различным опытом в ургентной нейрорадиологии и стажем менее 3 лет в первом варианте тестирования на выборке 1В

			а (TP) истинно положительные	б (FP) ложноположительные	с (FN) ложноотрицательные	д (TN) истинно отрицательные	Чувствительность, %	Специфичность, %	Точность, %	
Программа			44	27	6	23	88,00	46,00	67,00	
Специалисты РСЦ	1	до	49	9	1	41	98,00	82,00	90,00	
		1 этап	50	9	0	41	100,00	82,00	91,00	
	2	до	49	10	1	40	98,00	80,00	89,00	
		1 этап	49	12	1	38	98,00	76,00	87,00	
	3	до	50	12	0	38	100,00	76,00	88,00	
		1 этап	50	13	0	37	100,00	74,00	87,00	
	4	до	50	6	0	44	100,00	88,00	94,00	
		1 этап	50	7	0	43	100,00	86,00	93,00	
	Специалисты без опыта в РСЦ	5	до	47	8	3	42	94,00	84,00	89,00
			1 этап	49	7	1	43	98,00	86,00	92,00
6		до	48	31	2	19	96,00	38,00	67,00	
		1 этап	49	28	1	22	98,00	44,00	71,00	
7		до	36	17	14	33	72,00	66,00	69,00	
		1 этап	41	17	9	33	82,00	66,00	74,00	

Таблица 22 – Средние показатели диагностической эффективности врачей с различным опытом в ургентной нейрорадиологии и стажем менее 3 лет в первом варианте тестирования на выборке 1В

		а (TP) истинно положительные	б (FP) ложноположительные	с (FN) ложноотрицательные	д (TN) истинно отрицательные	Чувствительность, %	Специфичность, %	Точность, %
Программа		44	27	6	23	88,00	46,00	67,00
Специалисты РСЦ	до	49,50	9,25	0,50	40,75	99,00	81,50	90,25
	1 этап	49,75	10,25	0,25	39,75	99,50	79,50	89,50
Специалисты без опыта в РСЦ	до	43, (6)	18, (6)	6, (3)	31, (3)	67, (3)	62, (6)	75,00
	1 этап	46, (3)	17, (3)	3, (6)	32, (6)	92, (6)	65, (3)	79,00

Несмотря на показатели отдельных врачей, общие значения диагностической эффективности (точность, специфичность) среди врачей РСЦ в первом варианте тестирования снизились в большей степени за счет специфичности (с 81,5% до 79,5%) и увеличения количества ложноположительных ответов (с 9,25 до 10,25), способствуя возрастанию гипердиагностики в данной группе экспертов. Однако в группе специалистов без опыта в РСЦ отмечалось улучшение показателей эффективности (точность, чувствительность и специфичность) при первом варианте тестирования, в

основном за счет увеличения чувствительности (с 67,(3)% до 92, (6))% и сокращения ложноотрицательных ответов (с 6,(3) до 3,(6)), то есть со снижением уровня гиподиагностики. Данные результаты свидетельствуют о положительном влиянии искусственного интеллекта на диагностическую эффективность врачей без опыта при применении модели второго чтения.

Результаты первого совместного тестирования врачей и программы автоматического анализа инсульта на выборке КТ- изображений 2В представлены в таблице 23.

Таблица 23 – Результаты первоначальной оценки и первого совместного тестирования врачей с различным опытом в ургентной нейрорадиологии и стажем менее 3 лет с программой автоматического анализа

		Модель второго чтения			
		Первоначальная оценка		Первый вариант тестирования	
		≤ 7 и > 7	≤ 6 и > 6	≤ 7 и > 7	≤ 6 и > 6
Внутри группы	Специалисты РСЦ	0,593	0,541	0,598	0,633
	Специалисты СОП	0,234	0,260	0,365	0,304
Между всеми группами		0,425	0,409	0,530	0,464
Согласие с программой	Специалисты РСЦ	0,584	0,501	0,625	0,592
	Специалисты СОП	0,300	0,324	0,480	0,415
	Между всеми группами	0,440	0,411	0,556	0,480

При первоначальной оценке, как при дихотомическом делении результатов ≤ 7 и > 7 , так и ≤ 6 и > 6 , согласие специалистов РСЦ внутри группы было умеренным. При применении модели второго чтения не было выявлено значительных изменений показателей каппа Флейсса, что говорит об отсутствии значимого влияния системы поддержки врачебных решений на межэкспертную согласованность врачей из РСЦ. Аналогичная тенденция изменения согласованности наблюдалась и у врачей из СОП, показатели каппы Флейсса которых свидетельствовали об незначимом согласии между специалистами при первоначальной оценке и после первого тестирования на выборке 2В. Стоит отметить, что согласие с программой специалистов РСЦ при первоначальной оценке и при первом варианте тестирования с дихотомическим делением результатов ≤ 6 и > 6 уступало согласию друг с другом, однако показатели каппы Флейсса при дихотомическом делении результатов ≤ 7 и > 7 были выше при применении модели второго чтения. Также специалисты РСЦ при первоначальной оценке в 50,5% случаев выставили оценку ≤ 7 , при первом варианте тестирования этот показатель снизился до 42,5%. Аналогичным образом, врачи из РСЦ до предоставления интерпретации алгоритма искусственного интеллекта в 64,5 % случаев оценили распространенность ишемических изменений в бассейне средней мозговой артерии в ≤ 6 баллов по ASPECTS, однако при применении модели второго чтения этот показатель снизился до 57%. Данные результаты свидетельствуют об отсутствии значимых изменений согласованности между специалистами при внедрении модели второго чтения, но при этом их склонности снижать балл по ASPECTS после предоставления данных интерпретации алгоритма.

Также, мы оценили согласованность врачей при первоначальной оценке и первом тестировании на выборке 2В относительно каждой из зон по ASPECTS. Результаты представлены в таблице 24.

Таблица 24 – Показатели межэкспертного согласия среди врачей с различным опытом в ургентной нейрорадиологии и стажем менее 3 лет при первоначальной оценке областей ASPECTS и после первого варианта тестирования на выборке 2В КТ- изображений

	При первоначальной оценке						При втором варианте тестирования													
	С	IC	L	I	M1	M2	M3	M4	M5	M6	С	IC	L	I	M1	M2	M3	M4	M5	M6
Внутри группы																				
Специалисты РСЦ	0,520	0,165	0,480	0,504	0,317	0,525	0,478	0,652	0,518	0,472	0,648	0,227	0,557	0,643	0,393	0,631	0,654	0,696	0,534	0,493
Специалисты без опыта в РСЦ	0,520	0,070	0,141	0,109	0,250	0,010	0,138	0,183	0,134	0,157	0,534	0,041	0,400	0,217	0,410	0,294	0,442	0,428	0,732	0,495
Между всеми группами	0,497	0,090	0,321	0,238	0,312	0,281	0,325	0,470	0,335	0,340	0,566	0,204	0,443	0,379	0,428	0,466	0,549	0,594	0,523	0,483
Согласие с программой																				
Специалисты РСЦ + программа	0,532	0,191	0,440	0,431	0,335	0,484	0,480	0,632	0,516	0,472	0,389	0,091	0,455	0,657	0,297	0,167	0,294	0,500	0,420	0,269
Специалисты без опыта + программа	0,458	0,013	0,265	0,158	0,269	0,057	0,192	0,225	0,187	0,604	0,566	0,175	0,519	0,342	0,491	0,412	0,480	0,483	0,673	0,567
Между всеми группами и программой	0,496	0,112	0,337	0,248	0,317	0,284	0,342	0,467	0,350	0,371	0,396	0,114	0,409	0,402	0,347	0,415	0,322	0,475	0,446	0,325

Среди врачей из РСЦ при первом тестировании на выборке 2В согласие относительно каждой из областей ASPECTS выражено не изменилось, за исключением зоны М3 (от умеренного до существенного согласия). Более динамичная картина наблюдалась среди специалистов из СОП: наиболее заметно согласованность повысилась относительно зон М5 (от слабого к существенному), L, М3, М4 (от слабого к умеренному согласию).

Стоит отметить общее повышение значений каппа Флейсса относительно областей ASPECTS базально- ядерной области и островковой доли (по сравнению с “полушарными” зонами) после первого этапа тестирования на 2В выборке у врачей из РСЦ, однако в диапазоне умеренного согласия. Общая согласованность между врачами СОП в большей степени изменилась относительно “полушарных” областей (М1-М6) от слабой до умеренной.

При втором варианте тестирования, проведенного через 2 месяца, оценивалось влияние модели параллельного чтения на выборках 1С и 2С. Врачам сразу предоставлялись данные интерпретации КТ- изображений алгоритмом, после чего они могли как согласиться с системой автоматического анализа, так и предложить свой вариант оценки.

Результаты второго тестирования совместной выборок 1С и 2С КТ-изображений головного мозга врачами со стажем менее 3 лет и алгоритмом искусственного интеллекта представлены в таблицах 25 и 26.

Таблица 25 – Показатели диагностической эффективности врачей с различным опытом в ургентной нейрорадиологии и стажем менее 3 лет при втором тестировании на выборке 1С

			а (ТР) истинно положительные	б (FP) ложноположительные	с (FN) ложноотрицательные	д (TN) истинно отрицательные	Чувствительность, %	Специфичность, %	Точность, %
Программа			44	27	6	23	88,00	46,00	67,00
Специалисты РСЦ	1	до	49	9	1	41	98,00	82,00	90,00
		2 этап	50	4	0	46	100,00	92,00	96,00
	2	до	49	10	1	40	98,00	80,00	89,00
		2 этап	50	6	0	44	100,00	88,00	94,00
	3	до	50	12	0	38	100,00	76,00	88,00
		2 этап	50	5	0	45	100,00	90,00	95,00
	4	до	50	6	0	44	100,00	88,00	94,00
		2 этап	49	7	1	43	98,00	86,00	92,00
Специалисты без опыта в РСЦ	5	до	47	8	3	42	94,00	84,00	89,00
		2 этап	48	3	2	47	96,00	94,00	95,00
	6	до	48	31	2	19	96,00	38,00	67,00
		2 этап	48	16	2	34	96,00	68,00	82,00
	7	до	36	17	14	33	72,00	66,00	69,00
		2 этап	48	8	2	42	96,00	84,00	90,00

Таблица 26 – Средние показатели диагностической эффективности врачей с различным опытом в ургентной нейрорадиологии и стажем менее 3 лет при втором тестировании на выборке 1С

		a (TP) истинно положительные	b (FP) ложноположительные	c (FN) ложноотрицательные	d (TN) истинно отрицательные	Чувствительность, %	Специфичность, %	Точность, %
Программа		44,00	27,00	6,00	23,00	88,00	46,0	67,0
Специалисты РСЦ	до	49,50	9,25	0,50	40,75	99,00	81,50	90,25
	2 этап	49,75	5,50	0,25	44,50	99,50	89,00	94,25
Специалисты без опыта в РСЦ	до	43, (6)	18, (6)	6, (3)	31, (3)	67, (3)	62, (6)	75,00
	2 этап	48,00	9,00	2,00	41,00	96,00	82,00	89,00

Противоположно первому тестированию на выборке 1В, при втором варианте (на выборке 1С) средние показатели диагностической эффективности (точность, чувствительность и специфичность) врачей РСЦ увеличились в большей степени за счет точности (с 90,2% до 94,25%) и специфичности (с 81,5%

до 89%) с сокращением числа ложноположительных (с 18,(6) до 9) и ложноотрицательных (с 6,(3) до 2) ответов и снижением уровня гипо- и гипердиагностики. Подобная тенденция наблюдалась также среди специалистов без опыта работы в РСЦ. В данной группе специалистов отмечалась выраженная положительная динамика средних значений эффективности с нарастанием значений чувствительности (с 67,(3)% до 96%), специфичности (с 62,(6)% до 82%) и точности (с 75% до 89%), с сокращением числа ложноположительных (с 18,(6) до 9) и ложноотрицательных (с 6,(3) до 2) ответов. Результаты второго тестирования на выборке 1С свидетельствуют об улучшении интерпретации КТ-изображений пациентов с подозрением на ОНМК как среди врачей с опытом в РСЦ, так и без него при применении модели параллельного чтения.

Результаты второго тестирования на выборке 2С представлены в таблице 27.

Таблица 27 – Результаты первоначальной оценки и второго совместного тестирования врачей с различным опытом в ургентной нейрорадиологии и стажем менее 3 лет с программой автоматического анализа

		Модель параллельного чтения			
		Первоначальная оценка		Второй вариант тестирования	
		≤ 7 и > 7	≤ 6 и > 6	≤ 7 и > 7	≤ 6 и > 6
Внутри группы	Специалисты РСЦ	0,593	0,541	0,734	0,739
	Специалисты СОП	0,234	0,260	0,573	0,600

Между всеми группами		0,425	0,409	0,669	0,657
Согласие с программой	Специалисты РСЦ	0,584	0,501	0,734	0,743
	Специалисты СОП	0,300	0,324	0,658	0,646
	Между всеми группами	0,440	0,411	0,687	0,673

При первоначальной оценке (производилась при первом варианте тестирования), согласие специалистов РСЦ внутри группы было умеренным как при дихотомическом делении результатов ≤ 7 и > 7 , так и ≤ 6 и > 6 . При применении модели параллельного чтения было выявлено увеличение межэкспертной согласованности до существенной. Аналогичная тенденция изменения показателей согласия наблюдалась и у врачей из СОП, значения каппы Флейсса которых свидетельствовали о незначимом согласии между специалистами при первоначальной оценке, и повышении согласованности до умеренной (практически существенной) при втором тестировании. Стоит заметить, что согласие с программой специалистов РСЦ при первоначальной оценке с дихотомическим делением результатов ≤ 6 и > 6 и ≤ 7 и > 7 уступало согласию друг с другом, однако показатели каппы Флейсса при дихотомическом делении результатов ≤ 6 и > 6 и ≤ 7 и > 7 стали выше при применении модели параллельного чтения, что говорит о том, что специалисты более доверяли программе на данном этапе, чем при применении модели второго чтения. Также специалисты РСЦ при первоначальной оценке в 50,5% случаев выставили оценку ≤ 7 , при втором варианте тестирования на выборке 2С этот показатель снизился до 39% (в сравнении с первым вариантом тестирования - 42,5%). Врачи без опыта оценки инсульта (не из РСЦ) при первоначальной интерпретации выставили балл

≤ 7 в 68% случаев, на втором варианте тестирования на выборке 2С этот процент снизился до 53 (при первом этапе до 55,5%). Аналогично, врачи из РСЦ до предоставления интерпретации алгоритма искусственного интеллекта в 64,5 % случаев оценили распространенность ишемических изменений в бассейне средней мозговой артерии в ≤ 6 баллов по ASPECTS, при применении модели параллельного чтения этот показатель снизился до 52,5% (до 57% при применении модели второго чтения). Их коллеги без опыта в ургентной нейрорадиологии при первичной оценке в 78% случаев выставили ≤ 6 балл по ASPECTS, при втором тестировании в 63% (при первом в 65% случаев). Данные результаты свидетельствуют, что при увеличении согласия между врачами, специалисты склонны более выражено снижать балл по ASPECTS в случае применения модели параллельного чтения.

Мы оценили согласованность врачей при первоначальной оценке и втором тестировании на выборке 2С относительно каждой из зон по ASPECTS. Результаты представлены в таблице 28.

**Таблица 28 – Показатели межэкспертного согласия среди врачей с различным опытом в ургентной
нейрорадиологии и стажем менее 3 лет при первоначальной оценке областей ASPECTS и после второго
варианта тестирования на выборке 2С КТ- изображений**

	При первоначальной оценке						При втором варианте тестирования													
	С	IC	L	I	M1	M2	M3	M4	M5	M6	C	IC	L	I	M1	M2	M3	M4	M5	M6
Внутри группы																				
Специалисты РСЦ	0,520	0,165	0,480	0,504	0,317	0,525	0,478	0,652	0,518	0,472	0,796	0,557	0,567	0,528	0,571	0,668	0,564	0,619	0,640	0,735
Специалисты без опыта в РСЦ	0,520	0,070	0,141	0,109	0,250	0,010	0,138	0,183	0,134	0,157	0,640	0,747	0,457	0,365	0,547	0,606	0,478	0,672	0,622	0,525
Между всеми группами	0,497	0,090	0,321	0,238	0,312	0,281	0,325	0,470	0,335	0,340	0,731	0,587	0,551	0,449	0,550	0,646	0,506	0,650	0,653	0,592
Согласие с программой																				
Специалисты РСЦ + программа	0,532	0,191	0,440	0,431	0,335	0,484	0,480	0,632	0,516	0,472	0,815	0,614	0,605	0,535	0,586	0,683	0,592	0,639	0,664	0,674
Специалисты без опыта + программа	0,458	0,013	0,265	0,158	0,269	0,057	0,192	0,225	0,187	0,604	0,719	0,756	0,577	0,454	0,578	0,693	0,535	0,707	0,672	0,579
Между всеми группами и программой	0,496	0,112	0,337	0,248	0,317	0,284	0,342	0,467	0,350	0,371	0,754	0,623	0,583	0,476	0,565	0,669	0,533	0,662	0,667	0,595

Среди врачей из РСЦ при втором тестировании на выборке 2С согласие относительно каждой из областей ASPECTS значительно изменилось в отношении зон С, М2, М5, М6 (от умеренного до существенного согласия) и IC (от слабого до умеренного согласия). Сходная картина наблюдалась среди специалистов из СОП: согласованность выражено повысилась относительно зон IC, М2, М4, М5, М6 (от слабого к существенному), и L (от слабого к умеренному согласию).

Отмечалось равномерное общее повышение значений каппа Флейсса относительно областей ASPECTS (как базально- ядерной области и островковой доли, так и “полушарных” зон) после второго тестирования на 2С выборке у всех врачей от умеренного до существенного согласия.

По результатам пятого этапа диссертационного исследования можно сделать вывод, что модель второго чтения имеет отрицательное влияние на показатели врачей со стажем до 3 лет и опытом в оценке инсульта на КТ-изображениях. Также второе чтение менее эффективно улучшает показатели врачей без опыта в РСЦ и стажем до 3 лет. В то время, как применение модели параллельного чтения при внедрении алгоритма искусственного интеллекта в диагностический процесс врачей со стажем до 3 лет оказывает положительное влияние в сторону увеличения показателей диагностической эффективности выявления ОНМК в бассейне средней мозговой артерии и их межэкспертной согласованности при оценке по ASPECTS. Необходимо особо выделить отсутствие суммационного снижения диагностической эффективности при комплементарной оценке врачами и программным продуктом с тождественным положительным эффектом, в виде снижения уровня гипердиагностики у врачей.

ЗАКЛЮЧЕНИЕ

На сегодняшний день в диагностике ишемического инсульта нейровизуализация занимает ключевую позицию. С учетом ряда причин (время проведения исследования, отсутствия абсолютных противопоказаний, доступность), КТ является ведущим методом нейровизуализации в диагностике ОНМК, призванным первоначально разрешить две основные задачи: исключение ишемического поражения головного мозга и/или внутримозгового кровоизлияния. Распознавание признаков тромбоза магистральных артерий головного мозга и ишемии на ранних стадиях является сложной диагностической задачей, в особенности для рентгенологов с небольшим опытом работы. В условиях ежегодного обновления кадрового состава больниц молодыми специалистами, наличия периода их адаптации, а также нехватки лучевых диагностов для оперативной интерпретации лучевых исследований (в т.ч. КТ) в субъектах Российской Федерации, вполне вероятно закономерное повышение уровня вариативности интерпретации лучевых изображений врачами, и, как следствие, снижение качества оказания медицинской помощи.

С учетом того, что бассейн кровоснабжения средней мозговой артерии является наиболее часто поражаемым при ОНМК [53], с целью сформировать единый подход к диагностике инсульта в данном бассейне, в 2000 г. была разработана шкала ASPECTS (Alberta stroke program early CT score). Данная 10 – бальная система является, по мнению многих исследователей, надежным диагностическим методом. Однако ее использование имеет и ряд ограничений, в том числе высокую вариативность межэкспертного трактования распространенности ишемических изменений, которая может отразиться на дальнейшем ходе лечения пациента [3, 75]. Для частичного решения проблем субъективности оценки по шкале ASPECTS предлагается внедрение алгоритмов искусственного интеллекта в качестве систем поддержки врачебных решений. На сегодняшний день применение таких систем сосредоточено на анализе

бесконтрастных КТ, КТ- ангиографии и КТ- перфузии. Они направлены на определение балла по шкале ASPECTS, количественную оценку ядра инсульта, пенумбры, статуса коллатерального кровотока и локализации артериальной окклюзии в автоматическом режиме [113]. Применение таких алгоритмов предполагает более эффективное обнаружение ишемических изменений, сокращение числа случаев гипо- и гипердиагностики и снижение межэкспертной вариативности при оценке КТ- изображений пациентов, нуждающихся в экстренной медицинской помощи [3, 49, 113].

Для оценки влияния систем автоматического анализа на процесс принятия врачебных решений при диагностике ишемического инсульта в бассейне средней мозговой артерии методом компьютерной томографии было проведено исследование, включающее несколько этапов.

После изучения на первом этапе отечественной и зарубежной литературы, посвященной данной проблеме, на втором была сформирована и зарегистрирована база данных [19] на основе результатов исследований методом компьютерной томографии (КТ) 150 пациентов с клинической картиной ОНМК в бассейне средней мозговой артерии. В дальнейшем на основе данной базы были созданы две выборки КТ изображений для тестирования врачей и систем автоматического анализа. Первая выборка включала 50 КТ- исследований пациентов с подтвержденным ишемическим инсультом, и 50 КТ- исследований без патологических изменений по данным бесконтрастной КТ, КТ- ангиографии и КТ- перфузии. Вторая выборка содержала компьютерные томограммы только тех пациентов, у которых ОНМК по ишемическому типу был установлен, в том числе с помощью методов лучевой диагностики.

На третьем этапе были протестированы двадцать один рентгенолог с различным стажем и опытом в ургентной нейрорадиологии в первой части исследования, и пятнадцать рентгенологов из РСЦ во второй части. Первая выборка КТ- изображений была использована для определения диагностической эффективности специалистов, а также их согласованности в выявлении ранних

ишемических изменений, вторая выборка - для оценки согласия в отношении дихотомического деления балла по ASPECTS (объема РИИ в бассейне средней мозговой артерии). Выявлено влияние стажа работы в РСЦ врачей ургентной нейрорадиологии как на диагностическую эффективность, так и на согласованность в отношении большинства КТ-признаков ишемического инсульта. Точность специалистов с экспертной квалификацией (опытом оценки ишемического инсульта и стажем более 8 лет) составила 93,8%, в то время как у их коллег из РСЦ со стажем от 3 до 8 лет - 91,3%, со стажем до 3 лет - 88,3%. При этом стаж не оказал влияния на диагностическую эффективность врачей, не имеющих релевантного опыта в оценке инсульта (т.е. работающих в стационарах, не относящихся к РСЦ): у специалистов со стажем более 8 лет точность составила 82,3%, со стажем от 3 до 8 лет 85,0%, со стажем менее 3 лет - 75%. Также было выявлено, что при оценке КТ-исследований пациентов врачи независимо от стажа и опыта работы с ургентной патологией были склонны к гипердиагностике. Данная тенденция была более выражена среди врачей СОП со стажем менее 3 лет. Наиболее высокоспецифичным и воспроизводимым КТ-признаком ишемического инсульта стал признак гиперденсной СМА – все эксперты внутри своих групп показали наибольший уровень согласия (от 0,095 до 0,75). Самым неоднозначным признаком ишемического инсульта являлась сглаженность корковых борозд ввиду низкой согласованности среди врачей- рентгенологов при его трактовке (- 0,020 (случайное согласие) до 0,689). Признаки снижения плотности вещества головного мозга и нарушения дифференцировки коры имели умеренную и существенную согласованность среди врачей РСЦ (от 0,493 до 0,671). Далее была определена межэкспертная согласованность среди врачей РСЦ с различным стажем при определении распространенности ишемических изменений по шкале ASPECTS. Была выявлена высокая оценочная вариативность с низким уровнем согласия между специалистами вне зависимости от дихотомического деления – 6 и 7 (показания к тромбэкстракции и тромболизису) и их опыта (в среднем 0,391 (для ASPECTS 6) и 0,376 (для ASPECTS 7) - слабая согласованность). С учетом

умеренного и существенного согласия между врачами РСЦ в отношении признаков ишемического инсульта и высокой вариативности при его оценке по ASPECTS, данная шкала в нашем исследовании показала невысокие цифры воспроизводимости, что подтверждает наличие субъективности оценки по ASPECTS [4, 5].

Стоит отметить, что в зарубежной литературе отводится существенная роль исследованиям, направленным на определение межэкспертного согласия, поскольку именно стандартизированная оценка является залогом успешной диагностики ишемического инсульта в острейшей стадии [4]. В исследовании Farzin В. и соавт. [52] был изучен уровень межэкспертного и внутриэкспертного согласия при оценке по ASPECTS пятнадцатью экспертами с различной специализацией (неврологи, рентгенэндоваскулярные хирурги и нейрорадиологи) и стажем работы (преимущественно более 10 лет). Даже при дихотомическом делении (0–5/5–10 баллов) межэкспертная согласованность не достигла существенного уровня (0,561 каппа статистика), это значит, что как минимум 5 из 15 экспертов вынесут разное заключение в 15 % случаев. Внутриэкспертное согласие (согласие эксперта с собственной оценкой при первом и втором просмотре) варьировалось от 0,599 до 0,943 [4, 52]. Также Farzin В. и соавт. [52] проанализировали статьи, посвященные вариативности согласия при оценке по ASPECTS, опубликованные с 2000 по 2015 г. Методологии проанализированных исследований отличались друг от друга по нескольким характеристикам, в том числе была ли врачам предоставлена клиническая информация о пациенте, время, отведенное для оценки, доступ ко всем срезам КТ-исследования, возможность экспертов устанавливать свои собственные параметры окна. Результаты этого обзора текущего состояния воспроизводимости ASPECTS отражают широкую степень вариабельности межэкспертного согласия. Например, коэффициенты межэкспертного согласия (каппа Флейсса) в описанных исследованиях варьировались от 0,26 до 0,97 для ASPECTS без дихотомического деления и от 0,16 до 0,93 для ASPECTS с дихотомическим делением (≤ 6 и > 6) и (≤ 7 и > 7) [4,

52]. Несмотря на данную разнородность, межэкспертная согласованность при дихотомическом делении (вне зависимости [≤ 6 и > 6] или [≤ 7 и > 7]) относительно ASPECTS среди специалистов в литературе описывается как существенная [72], хотя в некоторых исследованиях сообщается об умеренном согласии [75].

В настоящем диссертационном исследовании мы также получили достаточно разнородные результаты, однако они в среднем не достигали значений существенного согласия между специалистами. Также мы получили обратно пропорциональную зависимость уровня межэкспертной согласованности от стажа работы специалиста (хотя в иностранной литературе данная зависимость прямо пропорциональна). Вероятно, это обусловлено тем, что несмотря на множественные пересмотры, в стандартизации шкалы существуют очевидные «слабые» места: отсутствует унифицированный подход к описанию характеристик РИИ, области ASPECTS не имеют точных границ, при внедрении шкалы в нейрорадиологическую практику в РФ не был принят единообразный и структурный подход к ее оценке, ввиду существовавших ограничений ее применения [4].

На четвертом этапе для изучения возможности применения систем автоматического применения при выявлении ишемического инсульта в бассейне средней мозговой артерии была проведена аналитическая валидация трех программ (А, В, С) для анализа компьютерных томограмм головного мозга. Первоначально оценка диагностической эффективности проводилась на 1 выборке изображений. Согласно клиническим рекомендациям (80%) по испытаниям программного обеспечения на основе технологий искусственного интеллекта [15], ни одна из задействованных систем не может быть допущена к клинической валидации со значениями точности в диапазоне от 67% до 75%. Программные продукты В и С продемонстрировали высокие показатели специфичности (0,9 и 0,8) и невысокие показатели чувствительности (0,6 и 0,55), что свидетельствует о редких случаях гипердиагностики и частых случаях

гиподиагностики. Для более достоверного понимания диагностических возможностей данных программных продуктов следует продолжать клинические испытания как методом аналитической валидации на различных выборках, так и методом клинической валидации. Однако программа А показала результаты, позволяющие судить о ее низком уровне гиподиагностики патологических изменений (чувствительность 88 истинно положительные ответы - 44 из 100, ложноположительные ответы 27 из 100). У врачей со стажем менее 3 лет и без опыта работы в РСЦ, по сравнению с программой А, ложноотрицательных ответов было больше в 3,5 раза, а истинно положительных ответов меньше в 1,5 раза. Относительно этих показателей врачи со стажем до 3 лет из РСЦ также несколько уступали программе А. С учетом потенциального положительного эффекта на молодых специалистов со стажем до 3 лет вне зависимости от опыта в оценке инсульта, в виде снижения случаев пропуска патологии, данный алгоритм искусственного интеллекта был выбран для дальнейшего исследования и протестирован на 2 выборке КТ-изображений [2].

Следует отметить, что показатель точности (0,80), согласно рекомендациям, позволяющий допустить программу до клинической валидации, в данном случае дискутабелен. Врачи РСЦ вне зависимости от стажа достигли более высоких показателей диагностической эффективности, чем уровень обозначенного параметра в клинических рекомендациях [2, 15].

К примеру, в исследовании П.В. Гаврилова и соавт. [8], посвященному выявлению округлых образований в легких на рентгенограммах в зависимости от стажа врача-рентгенолога и его опыта в работе в торакальной радиологии, точность специалистов без опыта составила 73,0%, а с опытом достигла 77,0%. Показатели чувствительности и специфичности в выявлении округлых образований легких также были несколько выше (81,0%; 75,0%), чем у врачей без опыта (75,0%; 72,0%), однако существенно не различались. При подобных результатах специалистов показатель, формально допускающий тестируемую программу к дальнейшей клинической валидации, представляется актуальным.

Однако по результатам тестирования врачей с различным стажем и опытом в диагностике ишемического инсульта, данный показатель релевантен только для врачей без опыта в ургентной медицине и со стажем менее 3 лет.

Также стоит учитывать наличие трудно диагностируемой патологии. К примеру, в диссертационном исследовании А. А. Мелдо [16], посвященном разработке и внедрению системы искусственного интеллекта в лучевой диагностике очаговых образований легких, выборки КТ- исследований, представленных для оценки врачей были разделены на трудно интерпретируемые случаи («типичный рак легкого», «не типичный рак легкого» и «не рак легкого»), и случаи с преимущественно типичными паттернами визуализации («типичный рак легкого», «не рак легкого»). Точность пяти врачей с различным (более 5 лет и менее 5 лет) стажем в диагностике онкологических заболеваний при интерпретации первой выборки составила от 51,0% до 71,1%, при оценке второй выборки от 82,0% до 96,0%. С учетом такого выраженного диапазона значений точности в исследовании, требуется дифференциация показателя, формально допускающего тестируемую программу к дальнейшей клинической валидации, в группе данной патологии.

Вероятно, существует необходимость учитывать текущие показатели эффективности врачей при диагностике различных патологических процессов для определения критериев допуска алгоритмов к клинической валидации. В противном случае, внедрение систем автоматического анализа может привести к отрицательной динамике точности выявления патологических изменений за счет интерференции в отношении специалистов без опыта, либо снизить доверие к программе за счет генерируемых очевидно неправильных результатов [2].

На пятом этапе исследования были изучены возможные варианты внедрения систем автоматического анализа компьютерно- томографических изображений в клиническую практику рентгенолога посредством проведения тестирования с моделированием двух различных ситуаций взаимодействия врача и программы А. Были применены две модели внедрения алгоритма

искусственного интеллекта в качестве системы поддержки врачебных решений: модель первого (параллельного) чтения и модель второго чтения. Модель параллельного чтения предполагала первоначальную оценку КТ-изображений алгоритмом ИИ, с последующей просмотром результатов диагностики системы автоматического анализа человеком. Применение модели второго чтения заключалась в просмотре исследований рентгенологом, с дальнейшей оценкой алгоритмом ИИ и заключительной интерпретацией данных врачом с возможным включением дополнительных результатов, обнаруженных ИИ. Тестирование проводилось на 1 и 2 выборке с определением влияния программы на диагностическую эффективность врачей со стажем до 3 лет с различным опытом в ургентной нейрорадиологии и их согласованность в отношении оценки по ASPECTS, соответственно. В тестировании приняли участие семь врачей со стажем работы менее 3 лет из второго этапа исследования. Врачи были распределены на две группы: специалисты ургентной медицины (РСЦ), и специалисты без опыта работы в РСЦ. При применении модели второго чтения, несмотря на показатели отдельных врачей, общие значения диагностической эффективности (точность, специфичность) среди врачей РСЦ снизились в большей степени за счет специфичности (с 81,5% до 79,5%) и увеличения количества ложноположительных ответов (с 9,25 до 10,25), способствуя возрастанию гипердиагностики в данной группе экспертов. Однако в группе специалистов без опыта в РСЦ отмечалось улучшение показателей эффективности (точность, чувствительность и специфичность) в основном за счет увеличения чувствительности (с 67,(3)% до 92, (6))% и сокращения ложноотрицательных ответов (с 6,(3) до 3,(6)), то есть со снижением уровня гиподиагностики. Данные результаты свидетельствуют о положительном влиянии искусственного интеллекта на диагностическую эффективность врачей без опыта при применении модели второго чтения.

При применении модели второго чтения не выявлено значительных изменений показателей каппа Флейсса (умеренное согласие) вне зависимости от

дихотомического деления баллов по ASPECTS (6 и 7), что говорит об отсутствии влияния системы поддержки врачебных решений на межэкспертную согласованность в отношении оценки распространенности ишемических изменений по ASPECTS. Согласованность врачей без опыта работы в РСЦ менялось с увеличением значений каппы Флейсса, но также в пределах незначимого согласия. Второе тестирование с применением модели параллельного чтения проводилось через 2 месяца. Средние показатели диагностической эффективности (точность, чувствительность и специфичность) врачей РСЦ увеличились в большей степени за счет точности (с 90,2% до 94,25%) и специфичности (с 81,5% до 89%) с сокращением числа ложноположительных (с 18,(6) до 9) и ложноотрицательных (с 6,(3) до 2) ответов, и снижением уровня гипо- и гипердиагностики. Подобная тенденция наблюдалась также среди специалистов без опыта работы в РСЦ. В данной группе специалистов, отмечалась выраженная положительная динамика средних значений эффективности с нарастанием значений чувствительности (с 67,(3)% до 96%), специфичности (с 62,(6) % до 82%) и точности (с 75% до 89%), с сокращением числа ложноположительных (с 18,(6) до 9) и ложноотрицательных (с 6,(3) до 2) ответов. Результаты применения модели параллельного чтения свидетельствуют об улучшении интерпретации КТ-изображений пациентов с подозрением на ОНМК при оценке алгоритма искусственного интеллекта как с врачами с опытом оценки инсульта, так и без него.

Также, при применении модели параллельного чтения среди специалистов РСЦ было выявлено увеличение межэкспертной согласованности от умеренной до существенной при оценке по ASPECTS с дихотомическим делением результатов ≤ 7 и > 7 , и ≤ 6 и > 6 . Аналогичная тенденция изменения согласованности наблюдалась и у врачей из СОП, показатели каппы Флейсса которых свидетельствовали об не значимом согласии между специалистами при первоначальной оценке и повышении согласованности до умеренной (практически существенной) при втором тестировании.

Результаты пятого этапа диссертационного исследования продемонстрировали выраженный положительный эффект применения модели параллельного чтения с увеличением диагностической эффективности врачей со стажем до 3 лет с повышением межэкспертного согласия при оценке по ASPECTS. Внедрение системы автоматического выявления ишемического инсульта в диагностический процесс менее опытных специалистов (со стажем работы менее 3 лет) из РСЦ позволяет им улучшить свою способность интерпретировать КТ-изображения, приближаясь к результативности специалистов с экспертной квалификацией (со стажем работы более 8 лет) за счет показателей чувствительности и точности.

Стоит отметить, что все включенные в исследование врачи как при применении модели параллельного, так и второго чтения были склонны к снижению балла по ASPECTS после совместной оценки с программой автоматического выявления инсульта. При первоначальном анализе специалисты РСЦ в 50,5% случаев выставили оценку ≤ 7 , при первом тестировании этот показатель снизился до 42,5%, при втором до 39%. Врачи без опыта оценки инсульта (не из РСЦ) при первоначальной интерпретации выставили балл ≤ 7 в 68% случаев, при первом тестировании этот процент снизился до 55,5, при втором - до 53%. Аналогично, врачи из РСЦ до предоставления интерпретации алгоритма искусственного интеллекта в 64,5 % случаев оценили распространенность ишемических изменений в бассейне средней мозговой артерии в ≤ 6 баллов по ASPECTS, при применении модели второго чтения этот показатель снизился до 57%, параллельного чтения – до 52,5%. Их коллеги без опыта в ургентной нейрорадиологии при первичной оценке в 78% случаев выставили ≤ 6 балл по ASPECTS, при первом тестировании в 65% случаев, при втором в 63%. Данные результаты свидетельствуют, что при увеличении согласия между врачами, специалисты склонны более выражено снижать балл по ASPECTS в случае применения модели параллельного чтения.

По данным иностранной литературы, зарубежными исследователями выявлены положительные результаты при взаимодействии врачей и систем автоматического анализа. К примеру, Brinjikji W. и соавт. [51] протестировали программное обеспечение e-ASPECTS совместно с двенадцатью рентгенологами и четырьмя неврологами. В исследование были включены нативные КТ-исследования 60 пациентов с подтвержденным ОНМК в бассейне средней мозговой артерии. Первоначально специалисты оценили бесконтрастные КТ-изображения головного мозга и выставили балл по ASPECTS, через два месяца врачи снова интерпретировали данные изображения, но уже совместно с программным обеспечением. Авторы выявили, что при использовании e-ASPECTS произошло значительное улучшение согласованности между наблюдателями. В исследовании не изучались особенности применения моделей внедрения систем поддержки врачебных решений, что не позволяет подтвердить результативность использования модели параллельного чтения.

В диссертационном исследовании Смольниковой У.А [20], посвященном возможностям систем автоматического анализа цифровых рентгенологических изображений в диагностике округлых образований в легких, несмотря на хорошие (в некоторых случаях превосходящие человека) результаты аналитической валидации систем автоматического анализа, при совместной интерпретации (модель параллельного чтения, согласно Obuchowski N. A. и соавт. [88]) рентгенограмм врачом и системами автоматического анализа, произошла суммация ошибок человека и программ, что привело к ухудшению параметров диагностической эффективности (снижение чувствительности с 83% до 56,7%, специфичности с 99%, до 93,9%). При применении модели второго чтения с первоначальной оценкой рентгенограмм врачом с последующим применением системы автоматического анализа и повторным принятием решения, выявлено снижение чувствительности до 66,7%, и увеличение специфичности - до 95% (на 10,6%). Стоит отметить, что в данном исследовании применение модели второго чтения было более целесообразно, поскольку улучшало чувствительность оценки

врачей, что отличается от результатов настоящего диссертационного исследования. Вероятно, стоит персонализировано подходить не только к выбору системы автоматического анализа, но и дополнительно оценивать особенности взаимодействия врача и искусственного интеллекта в зависимости от области его специализации.

Исходя из вышесказанного, несмотря на положительные результаты применения систем автоматического анализа, необходимы дальнейшие исследования по выработке их оптимального взаимодействию с врачами для улучшения показателей диагностической эффективности. Однако уже на этом этапе, возможно судить о перспективах внедрения алгоритмов искусственного интеллекта в клиническую практику врача.

ВЫВОДЫ

1. Выявлена прямая зависимость уровня диагностической эффективности в выявлении ишемических изменений в бассейне средней мозговой артерии с помощью компьютерной томографии от стажа врачей рентгенологов и наличия у них опыта работы в региональном сосудистом центре.
2. Выявлена существенная согласованность в отношении признаков гиперденсной средней мозговой артерии (0,684) и нарушения дифференцировки серого и белого вещества (0,629), а также умеренное согласие относительно признака сглаженности корковых борозд (0,525) и снижения КТ- плотности вещества головного мозга (0,529) среди врачей региональных сосудистых центров.
3. Несмотря на существенную согласованность в отношении отдельных КТ- признаков ишемического инсульта (в диапазоне от 0,529 до 0,684) и высокую диагностическую эффективность при определении ОНМК в бассейне средней мозговой артерии (точность в диапазоне от 88,3 до 93,8%), выявлены низкие показатели воспроизводимости шкалы ASPECTS (0,391 при дихотомическом делении баллов по ASPECTS с границей 6 и 0,376 при дихотомическом делении баллов по ASPECTS с границей 7) у врачей регионального сосудистого центра вне зависимости от стажа.
4. Параметры диагностической эффективности существующих в настоящее время систем автоматического анализа компьютерных томограмм (точность в диапазоне от 67,0 до 75,0%) при выявлении ишемического инсульта в бассейне средней мозговой артерии ниже по сравнению с результатами врачей регионального сосудистого центра вне зависимости от стажа (точность в диапазоне от 88,3 до 93,8%).
5. Оптимальным вариантом внедрения систем автоматического анализа с низким уровнем гиподиагностики и высоким риском гипердиагностики

является применение модели, в которой первоначально КТ- исследования оцениваются программным обеспечением с последующей их оценкой врачом и принятием решения с учетом результатов интерпретации алгоритма искусственного интеллекта. Такой подход позволяет увеличить показатели диагностической эффективности врачей со стажем менее 3 лет (точность от 90,25 до 94,25% среди врачей РСЦ, и точность от 75 до 89% среди врачей СОП), а также снизить вариативность межэкспертной оценки ишемических изменений по ASPECTS (от 0,409 до 0,657 при дихотомическом делении баллов по ASPECTS с границей 6 и от 0,425 до 0,669 при дихотомическом делении баллов по ASPECTS с границей 7).

ПРАКТИЧЕСКИЕ РЕКОМЕНДАЦИИ

1. Применение систем автоматического анализа позволяет увеличить показатели диагностической эффективности врачей рентгенологов со стажем до 3 лет в выявлении ишемического инсульта в бассейне средней мозговой артерии с помощью компьютерной томографии.
2. Применение систем автоматического анализа позволяет увеличить воспроизводимость шкалы ASPECTS среди врачей рентгенологов со стажем до 3 лет.
3. Существующие в настоящий момент программы автоматической оценки ишемических изменений различаются по показателям диагностической эффективности и не достигают значений, необходимых для дальнейшей клинической валидации.
4. Рекомендуется тестирование внедряемого алгоритма на независимой верифицированной выборке КТ-изображений
5. При тестировании систем автоматического анализа наибольшее значение имеет оценка параметров, указывающих на высокую чувствительность, что влияет на снижение риска гиподиагностики патологических изменений.
6. При внедрении систем автоматического анализа в диагностический процесс, целесообразно использовать модель, когда первоначально КТ-исследования оцениваются программным обеспечением с последующей их оценкой врачом и заключительным принятием решения с учетом результатов интерпретации алгоритма искусственного интеллекта.

ПЕРСПЕКТИВЫ ДАЛЬНЕЙШЕЙ РАЗРАБОТКИ ТЕМЫ

Своевременное выявление и лечение острого ишемического инсульта имеет решающее значение для снижения заболеваемости и смертности населения РФ. Применение ИИ в диагностике ишемического инсульта в бассейне средней мозговой артерии может предоставить широкие возможности для более точной, а в последствии и быстрой интерпретации лучевых изображений, что в итоге может улучшить качество оказания медицинской помощи и функциональные исходы заболевания.

Одними из наиболее перспективных направлений внедрения систем искусственного интеллекта, является создание платформ, инкорпорирующих программные продукты, направленные на распознавания ишемического инсульта методом компьютерной томографии, острых окклюзий брахицефальных и церебральных артерий, а также автоматическую оценку КТ- перфузии. Также необходимы дальнейшие исследования по выявлению ишемического инсульта в вертебрально- базиллярном бассейне с помощью искусственного интеллекта, а также создание прогностических моделей развития заболевания. Наиболее актуальными из таких моделей являются прогнозирование исхода ТЛТ, а также результатов других методов лечения.

Остается невыясненным вопрос, будет ли выявлена эффективность данной интеграции в клиническую практику? И как это повлияет на дальнейший суммационный эффект диагностических и терапевтических мероприятий, скажется на правовом секторе и юридической ответственности врачей и учреждений здравоохранения в спорных случаях, даже при наличии официальных руководств и рекомендаций по применению систем автоматического анализа.

Необходимы дальнейшие исследования, подтверждающие результативность методов искусственного интеллекта, чтобы обеспечить снижение вариативности интерпретаций лучевых изображений и более широкое использование в различных практических условиях. Для более полного понимания

диагностических возможностей программных продуктов искусственного интеллекта следует продолжать клинические испытания как методом аналитической валидации на различных выборках, так и методом клинической валидации.

СПИСОК СОКРАЩЕНИЙ

ВТТ – внутривенная тромболитическая терапия

ДВИ – диффузионно - взвешенные изображения

ИИ – искусственный интеллект

КТ – компьютерная томография

МРТ – магнитно - резонансная томография

ОНМК – острое нарушение мозгового кровообращения

ПСО – первичные сосудистые отделения

РИИ – ранние ишемические изменения

РСЦ – региональный сосудистый центр

РФ – Российская Федерация

СМА – средняя мозговая артерия

СОП – стационары общего профиля

ASPECTS (Alberta stroke program early CT score) – программа Alberta для оценки ранних КТ изменений при инсульте, 10-балльная шкала полуколичественной топографической оценки ишемических изменений, выявляемых при компьютерной томографии.

DICOM (Digital Imaging and Communications in Medicine) – медицинский отраслевой стандарт создания, хранения, передачи и визуализации цифровых медицинских изображений и документов обследованных пациентов

JPEG (Joint Photographic Experts Group) – один из растровых графических форматов, применяемый для хранения фотографий и подобных им изображений.

mRS (Modified Rankin scale) – шкала Рэнкина

NIHSS (National Institutes of Health Stroke Scale) – шкала инсульта национального института здоровья

PACS (Picture Archiving and Communication System) — это система управления клиническими данными, состоящая из нескольких технологий медицинской визуализации с открытым исходным кодом, которая может использоваться для

хранения, доступа, передачи и управления медицинскими изображениями и цифровыми отчетами в формате DICOM.

СЛОВАРЬ ТЕРМИНОВ

Аналитическая валидация – оценка корректности обработки входных данных программным обеспечением для создания надежных выходных данных; оценивается с применением эталонных наборов размеченных данных.

Гипердиагностика – ошибочное медицинское заключение о наличии у обследуемого болезни или ее осложнений, которые на самом деле отсутствуют либо выражены слабее, чем указано в заключении.

Гиподиагностика – ошибочное медицинское заключение об отсутствии у обследуемого болезни или ее осложнений, которые на самом деле присутствуют либо выражены сильнее, чем указано в заключении.

Глубокое обучение – совокупность методов машинного обучения (с учителем, с частичным привлечением учителя, без учителя, с подкреплением), основанных на обучении представлениям, а не специализированных алгоритмах под конкретные задачи.

Диагностическая эффективность – это параметр, характеризующий возможности данной тест-системы одновременно правильно определять положительные результаты как положительные, а отрицательные результаты — как отрицательные.

Искусственный интеллект – комплекс технологических решений, позволяющий имитировать когнитивные функции человека (включая самообучение и поиск решений без заранее заданного алгоритма) и получать при выполнении конкретных задач результаты, сопоставимые, как минимум, с результатами интеллектуальной деятельности человека.

Клиническая валидация – подтверждение достижения намеченной цели в целевой популяции в контексте лечебной работы благодаря использованию точных и надежных выходных данных.

Машинное обучение – класс методов искусственного интеллекта, характерной чертой которых является не прямое решение задачи, а обучение за счет применения решений множества сходных задач. Для построения таких методов используются средства математической статистики, численных методов, математического анализа, методов оптимизации, теории вероятностей, теории графов, различные техники работы с данными в цифровой форме.

Нейронная сеть – математическая модель, а также её программное или аппаратное воплощение, построенная по принципу организации и функционирования биологических нейронных сетей — сетей нервных клеток живого организма, с точки зрения машинного обучения, нейронная сеть представляет собой частный случай методов распознавания образов, дискриминантного анализа.

Система автоматического анализа изображений – система, позволяющая производить обнаружение, отслеживание и классификацию объектов, в частности обработку изображений в медицине, способствует получению информации из данных изображения для постановки медицинского диагноза пациентам.

СПИСОК ЛИТЕРАТУРЫ

1. Анализ эпидемиологических показателей повторных инсультов в регионах Российской Федерации (по итогам территориально-популяционного регистра 2009–2014 гг.) / Л. В. Стаховская, О. А. Клочихина, М. Д. Богатырева, С. А. Чугунова. – 2016. – № 18 (9). – С. 8–11
2. Андропова, П. Л. Диагностическая эффективность отдельных систем автоматического анализа КТ-изображений в выявлении ишемического инсульта в бассейне средней мозговой артерии / П. Л. Андропова, П. В. Гаврилов, П.А. Колесникова [и др.] // Сибирский журнал клинической и экспериментальной медицины. – 2023. – № 3(39). – С. 194-200.
3. Андропова, П. Л. Шкала ASPECTS: межэкспертное соглашение при использовании врачами-рентгенологами / П. Л. Андропова, П. В. Гаврилов, Ж. И. Савинцева // Лучевая диагностика и терапия. – 2022. – № 1(13). – С. 7-13.
4. Андропова, П. Л. Оценка межэкспертной согласованности врачей-рентгенологов в диагностике ишемического инсульта в бассейне средней мозговой артерии с помощью компьютерной томографии / П. Л. Андропова, П. В. Гаврилов, И. П. Казанцева [и др.] // Медицинская визуализация. – 2023. – Режим доступа: <https://doi.org/10.24835/1607-0763-1315>.
5. Андропова, П.Л. Оценка межэкспертной согласованности при использовании ASPECTS врачами ургентной нейрорадиологии с различным стажем / П. Л. Андропова, П. В. Гаврилов, И. П. Казанцева [и др.] // Радиология - практика. – 2022. – № 5(95). – С. 10-25.
6. Андропова, П.Л. Применение систем искусственного интеллекта в нейрорадиологии острого ишемического инсульта / П. Л. Андропова, П. В. Гаврилов, Ж. И. Савинцева [и др.] // Лучевая диагностика и терапия. – 2021. – № 2(12). – С. 30-35.

7. Васильев, А. Ю. Анализ данных лучевых методов исследования на основе принципов доказательной медицины/ Васильев А. Ю., Малый А. Ю., Серов Н. С. - Москва: ГЭОТАР-Медиа, 2008. – 32 с.
8. Гаврилов, П. В. Выявление округлых образований в легких при цифровой рентгенографии: роль опыта работы врача-рентгенолога / П. В. Гаврилов, А. Д. Ушков, У. А. Смольникова // Медицинский альянс. - 2019. - № 2. - С. 51–56
9. Дорожная карта развития “сквозной” цифровой технологии “Нейротехнологии и искусственный интеллект”. Режим доступа: <https://digital.gov.ru/ru/documents/6658>. Дата обращения 17 июня, 2020.
10. Инсульт: Клиническое руководство / М.Г. Хеннерици, Ж. Богуславски, Р.Л. Сакко. Пер. с англ. – Под общ.ред. чл.корр. РАМН Скворцовой В.И. 2е изд. М.: МЕДпрессинформ, 2008. – 224 с.
11. Инсульт: Руководство для врачей / Под ред. Л.В. Стаховской, С.В. Котова. — М.: ООО «Медицинское информационное агентство», 2013. —400 с.: ил.
12. Ишемический инсульт и транзиторная ишемическая атака у взрослых: клинические рекомендации / Р.Г. Акжигитов, Б.Г. Алекян, В.В. Алферова [и др.]. – М.: Министерство здравоохранения Российской Федерации, 2021. – 181 с.
13. Ишемический инсульт и транзиторная ишемическая атака у взрослых: клинические рекомендации / Всероссийское общество неврологов [и др.]. – М.: Министерство здравоохранения Российской Федерации, с изм. и допол. в ред. от 2022. – с. 214
14. Ишемический инсульт: злокачественный инфаркт в бассейне средней мозговой артерии: клинические рекомендации / В.В. Крылов, О.Н. Древаль, Р.С. Джинджихадзе [и др.]. – М.: Министерство здравоохранения Российской Федерации, 2015. – 17 с.
15. Клинические испытания программного обеспечения на основе интеллектуальных технологий (Лучевая Диагностика) / С. П. Морозов, А. В.

- Владзимирский, В. Г. Кляшторный [и др.] // Лучшие практики лучевой и инструментальной диагностики Выпуск 23. - 2019. - 34с.
16. Мелдо А. А. Разработка и внедрение системы искусственного интеллекта в лучевой диагностике очаговых образований в легких: дис. ... д-ра мед. наук; 3.1.25 / А.А. Мелдо. — Санкт-Петербург, 2022. — 235. Режим доступа: https://disser.spbu.ru/files/2021/disser_meldo.pdf (29.08.2023)
17. Национальный стандарт РФ ГОСТ Р 59921.0-2022 "Системы искусственного интеллекта в клинической медицине. Основные положения"
18. Национальный стандарт РФ ГОСТ Р 59921.5-2022 «Системы искусственного интеллекта в клинической медицине. Часть 5. Требования к структуре и порядку применения набора данных для обучения и тестирования алгоритмов»
19. Результаты компьютерно-томографических исследований пациентов с клинической картиной острого нарушения мозгового кровообращения в бассейне средней мозговой артерии / П. Л. Андропова, П. В. Гаврилов, Ж. И. Савинцева [и др.] // Свидетельство о регистрации базы данных RU 2022620850, 18.04.2022. Заявка № 2022620723 от 07.04.2022
20. Смольникова У.А. Возможности систем автоматического анализа цифровых рентгенологических изображений в диагностике округлых образований в легких: дис. ... канд. мед. наук: 3.1.25 / У.А. Смольникова. — Санкт-Петербург, 2022. — 281 с. Режим доступа: https://disser.spbu.ru/files/2022/disser_smolnikoiva.pdf (01.09.2022).
21. Указ Президента РФ от 10.10.2019 № 490 «О развитии искусственного интеллекта в Российской Федерации» (вместе с «Национальной стратегией развития искусственного интеллекта на период до 2030 года»). Режим доступа: http://www.consultant.ru/document/cons_doc_LAW_335184/. Дата обращения: 15.03.2022.
22. Фейгин, В. Инсульт: Клиническое руководство / В. Фейгин, Д. Виберс, Р. Браун. – М.: Бином, 2005. – 307 с.

23. Эпидемиология инсульта в России по результатам территориально-популяционного регистра (2009-2010) / Л. В. Стаховская, О. А. Ключихина, М. Д. Богатырева, В. В. Коваленко // Журнал неврологии и психиатрии. – 2013. – №5. – С. 4-10.
24. Advanced [18F] FDG and [11C] flumazenil PET analysis for individual outcome prediction after temporal lobe epilepsy surgery for hippocampal sclerosis / J. Yankam Njiwa, K. R. Gray, N. Costes [et al.] // NeuroImage: Clinical. – 2015. – Vol. 7. – P. 122-131.
25. Advances in Auto-Segmentation / C. E. Cardenas, J. Yang, B. M. Anderson [et al.] // Seminars in Radiation Oncology. – 2019. – Vol. 29. – № 3. – P. 185-197.
26. Agreement and Variability in the Interpretation of Early CT Changes in Stroke Patients Qualifying for Intravenous rtPA Therapy / J. C. Grotta, D. Chiu, M. Lu [et al.] // Stroke. – 1999. – Vol. 30. – № 8. – P. 1528-1533.
27. Alberta Stroke Program Early Computed Tomographic Scoring Performance in a Series of Patients Undergoing Computed Tomography and MRI: Reader Agreement, Modality Agreement, and Outcome Prediction / R. A. McTaggart, T. G. Jovin, M. G. Lansberg [et al.] // Stroke. – 2015. – Vol. 46. – № 2. – P. 407-412.
28. Alberta Stroke Program Early CT Scoring of CT Perfusion in Early Stroke Visualization and Assessment / R. I. Aviv, J. Mandelcorn, S. Chakraborty [et al.] // American Journal of Neuroradiology. – 2007. – Vol. 28. – № 10. – P. 1975-1980.
29. Andropova P. A. Computer-Aided Diagnosis System in the Diagnosis of Multiple Sclerosis / P. Andropova, D. Cheremisin, A. Meldo. – Text: electronic // Proceedings of International Scientific Conference on Telecommunications, Computing and Control: Smart Innovation, Systems and Technologies / eds. N. Voinov [et al.]. – Singapore: Springer Singapore, 2021. – Vol. 220. – P. 277-284. – URL: https://link.springer.com/10.1007/978-981-33-6632-9_24 (date accessed: 22.08.2023).

30. Application of Machine Learning to Automated Analysis of Cerebral Edema in Large Cohorts of Ischemic Stroke Patients / R. Dhar, Y. Chen, H. An, J.-M. Lee // *Frontiers in Neurology*. – 2018. – Vol. 9. – P. 687.
31. Artificial Intelligence and Acute Stroke Imaging / J. E. Soun, D. S. Chow, M. Nagamine [et al.] // *AJNR. American journal of neuroradiology*. – 2021. – Vol. 42. – № 1. – P. 2-11.
32. Artificial Intelligence and Machine Learning in Radiology: Opportunities, Challenges, Pitfalls, and Criteria for Success / J. H. Thrall, X. Li, Q. Li [et al.] // *Journal of the American College of Radiology*. – 2018. – Vol. 15. – № 3. – P. 504-508.
33. Artificial intelligence-driven ASPECTS for the detection of early stroke changes in non-contrast CT: a systematic review and meta-analysis / A. Adamou, E. T. Beltsios, A. Bania [et al.] // *Journal of NeuroInterventional Surgery*. – 2022. – P. jnis-2022-019447.
34. Assistance from Automated ASPECTS Software Improves Reader Performance / P. R. Delio, M. L. Wong, J. P. Tsai [et al.] // *Journal of Stroke and Cerebrovascular Diseases*. – 2021. – Vol. 30. – № 7. – P. 105829.
35. Automated quantification of cerebral edema following hemispheric infarction: Application of a machine-learning algorithm to evaluate CSF shifts on serial head CTs / Y. Chen, R. Dhar, L. Heitsch [et al.] // *NeuroImage: Clinical*. – 2016. – Vol. 12. – P. 673-680.
36. Butcher K. Acute Stroke Imaging Part I: Fundamentals / K. Butcher, D. Emery // *Canadian Journal of Neurological Sciences / Journal Canadien des Sciences Neurologiques*. – 2010. – Vol. 37. – № 1. – P. 4-16.
37. Can structural MRI aid in clinical classification? A machine learning study in two independent samples of patients with schizophrenia, bipolar disorder and healthy subjects / H. G. Schnack, M. Nieuwenhuis, N. E. M. Van Haren [et al.] // *NeuroImage*. – 2014. – Vol. 84. – P. 299-306.

38. Chen L. Fully automatic acute ischemic lesion segmentation in DWI using convolutional neural networks / L. Chen, P. Bentley, D. Rueckert // *NeuroImage: Clinical*. – 2017. – Vol. 15. – P. 633-643.
39. Classifying Acute Ischemic Stroke Onset Time using Deep Imaging Features / K. C. Ho, W. Speier, S. El-Saden, C. W. Arnold // *AMIA ... Annual Symposium proceedings. AMIA Symposium*. – 2017. – Vol. 2017. – P. 892-901.
40. Comparison of the performance between Frontier ASPECTS software and different levels of radiologists on assessing CT examinations of acute ischaemic stroke patients / L. Li, Y. Chen, Y. Bao [et al.] // *Clinical Radiology*. – 2020. – Vol. 75. – № 5. – P. 358-365.
41. Computer-assisted identification and quantification of multiple sclerosis lesions in MR imaging volumes in the brain / J. R. Mitchell, S. J. Karlik, D. H. Lee, A. Fenster // *Journal of Magnetic Resonance Imaging*. – 1994. – Vol. 4. – № 2. – P. 197-208.
42. Convolutional neural networks: an overview and application in radiology / R. Yamashita, M. Nishio, R. K. G. Do, K. Togashi // *Insights into Imaging*. – 2018. – Vol. 9. – № 4. – P. 611-629.
43. Crowdsourcing image analysis for plant phenomics to generate ground truth data for machine learning / N. Zhou, Z. D. Siegel, S. Zarecor [et al.] // *PLOS Computational Biology*. – 2018. – Vol. 14. – № 7. – P. e1006337.
44. Cruz J. A. Applications of machine learning in cancer prediction and prognosis / J. A. Cruz, D. S. Wishart // *Cancer Informatics*. – 2007. – Vol. 2. – P. 59-77.
45. CT interpretation by ASPECTS in hyperacute ischemic stroke predicting functional outcomes / W. Phuttharak, K. Sawanyawisuth, B. Sangpetngam, S. Tiamkao // *Japanese Journal of Radiology*. – 2013. – Vol. 31. – № 10. – P. 701-705.
46. CT Perfusion ASPECTS in the Evaluation of Acute Ischemic Stroke: Thrombolytic Therapy Perspective / N. Sillanpaa, J. T. Saarinen, H. Rusanen [et al.] // *Cerebrovascular Diseases Extra*. – 2011. – Vol. 1. – № 1. – P. 6-16.

47. CT visualization of intracranial arterial thromboembolism. / G. Gács, A. J. Fox, H. J. Barnett, F. Vinuela // *Stroke*. – 1983. – Vol. 14. – № 5. – P. 756-762.
48. Decoding post-stroke motor function from structural brain imaging / J. M. Rondina, M. Filippone, M. Girolami, N. S. Ward // *NeuroImage: Clinical*. – 2016. – Vol. 12. – P. 372-380.
49. Dermatologist-level classification of skin cancer with deep neural networks / A. Esteva, B. Kuprel, R. A. Novoa [et al.] // *Nature*. – 2017. – Vol. 542. – № 7639. – P. 115-118.
50. Differential Prognosis of Isolated Cortical Swelling and Hypoattenuation on CT in Acute Stroke / K. S. Butcher, S. B. Lee, M. W. Parsons [et al.] // *Stroke*. – 2007. – Vol. 38. – № 3. – P. 941-947.
51. e-ASPECTS software improves interobserver agreement and accuracy of interpretation of aspects score / W. Brinjikji, M. Abbasi, C. Arnold [et al.] // *Interventional Neuroradiology*. – 2021. – Vol. 27. – № 6. – P. 781-787.
52. Early CT changes in patients admitted for thrombectomy: Intrarater and interrater agreement / B. Farzin, R. Fahed, F. Guilbert [et al.] // *Neurology*. – 2016. – Vol. 87. – № 3. – P. 249-256.
53. Early Identification of Potentially Salvageable Tissue with MRI-Based Predictive Algorithms after Experimental Ischemic Stroke / M. J. Bouts, I. A. Tiebosch, A. Van Der Toorn [et al.] // *Journal of Cerebral Blood Flow & Metabolism*. – 2013. – Vol. 33. – № 7. – P. 1075-1082.
54. Early prediction of irreversible brain damage after ischemic stroke at CT / R. von Kummer, H. Bourquain, S. Bastianello [et al.] // *Radiology*. – 2001. – Vol. 219. – № 1. – P. 95-100.
55. Effect of baseline Alberta Stroke Program Early CT Score on safety and efficacy of intra-arterial treatment: a subgroup analysis of a randomised phase 3 trial (MR CLEAN) / A. J. Yoo, O. A. Berkhemer, P. S. S. Fransen [et al.] // *The Lancet. Neurology*. – 2016. – Vol. 15. – № 7. – P. 685-694.

56. Effects of Alteplase for Acute Stroke on the Distribution of Functional Outcomes: A Pooled Analysis of 9 Trials / K. R. Lees, J. Emberson, L. Blackwell [et al.] // *Stroke*. – 2016. – Vol. 47. – № 9. – P. 2373-2379.
57. Endovascular Treatment of Middle Cerebral Artery M2 Occlusion Strokes: Clinical and Procedural Predictors of Outcomes / M. Mokin, C. T. Primiani, Z. Ren [et al.] // *Neurosurgery*. – 2017. – Vol. 81. – № 5. – P. 795-802.
58. Extent of hypoattenuation on CT angiography source images predicts functional outcome in patients with basilar artery occlusion / V. Puetz, P. N. Sylaja, S. B. Coutts [et al.] // *Stroke*. – 2008. – Vol. 39. – № 9. – P. 2485-2490.
59. Feng G. C. Mistakes and How to Avoid Mistakes in Using Intercoder Reliability Indices / G. C. Feng // *Methodology*. – 2015. – Vol. 11. – № 1. – P. 13-22.
60. Fiebach J. B. Moderne Kernspintechiken beim Schlaganfall / J. B. Fiebach, P. D. Schellinger // *Der Radiologe*. – 2003. – Vol. 43. – № 3. – P. 251-264.
61. Gillies R. J. Radiomics: Images Are More than Pictures, They Are Data / R. J. Gillies, P. E. Kinahan, H. Hricak // *Radiology*. – 2016. – Vol. 278. – Radiomics. – № 2. – P. 563-577.
62. Global, regional, and national burden of stroke, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016 / C. O. Johnson, M. Nguyen, G. A. Roth [et al.] // *The Lancet Neurology*. – 2019. – Vol. 18. – № 5. – P. 439-458.
63. Guidelines for the Early Management of Adults With Ischemic Stroke: A Guideline From the American Heart Association/ American Stroke Association Stroke Council, Clinical Cardiology Council, Cardiovascular Radiology and Intervention Council, and the Atherosclerotic Peripheral Vascular Disease and Quality of Care Outcomes in Research Interdisciplinary Working Groups: The American Academy of Neurology affirms the value of this guideline as an educational tool for neurologists / H. P. Adams, G. Del Zoppo, M. J. Alberts [et al.] // *Stroke*. – 2007. – Vol. 38. – № 5. – P. 1655-1711.

64. Hallgren K. A. Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial / K. A. Hallgren // *Tutorials in Quantitative Methods for Psychology*. – 2012. – Vol. 8. – № 1. – P. 23-34.
65. Ho M.-L. Cerebral Edema / M.-L. Ho, R. Rojas, R. L. Eisenberg // *American Journal of Roentgenology*. – 2012. – Vol. 199. – № 3. – P. W258-W273.
66. Hsu K.-C. Backward somersault as a cause of childhood stroke: a case report of isolated middle cerebral artery dissection in an adolescent boy / K.-C. Hsu, H.-W. Kao, S.-J. Chen // *The American Journal of Emergency Medicine*. – 2008. – Vol. 26. – № 4. – P. 519.e3-519.e5.
67. Huang S. Quantitative prediction of acute ischemic tissue fate using support vector machine / S. Huang, Q. Shen, T. Q. Duong // *Brain Research*. – 2011. – Vol. 1405. – P. 77-84.
68. Hyperdense Sylvian Fissure MCA 'Dot' Sign: A CT Marker of Acute Ischemia / P. A. Barber, A. M. Demchuk, M. E. Hudon [et al.] // *Stroke*. – 2001. – Vol. 32. – № 1. – P. 84-88.
69. Hyperacute stroke: experience essential when reading unenhanced CT scans / J. H. W. Pexman, M. D. Hill, A. M. Buchan [et al.] // *AJNR. American journal of neuroradiology*. – 2004. – Vol. 25. – № 3. – P. 516; author reply 516-518.
70. Identifying spatial imaging biomarkers of glioblastoma multiforme for survival group prediction: GBM Survival Group Prediction / M. Zhou, B. Chaudhury, L. O. Hall [et al.] // *Journal of Magnetic Resonance Imaging*. – 2017. – Vol. 46. – № 1. – P. 115-123.
71. Impact of Window Setting Optimization on Accuracy of Computed Tomography and Computed Tomography Angiography Source Image-based Alberta Stroke Program Early Computed Tomography Score / E. M. Arsava, J. T. Saarinen, A. Unal [et al.] // *Journal of Stroke and Cerebrovascular Diseases*. – 2014. – Vol. 23. – № 1. – P. 12-16.
72. Interobserver Variation of ASPECTS in Real Time / S. B. Coutts, A. M. Demchuk, P. A. Barber [et al.]. – Text: electronic // *Stroke*. – 2004. – Vol. 35. – № 5. – URL:

<https://www.ahajournals.org/doi/10.1161/01.STR.0000127082.19473.45> (date accessed: 22.08.2023).

73. Ischemic Brain Tissue Water Content: CT Monitoring during Middle Cerebral Artery Occlusion and Reperfusion in Rats / I. Dzialowski, E. Klotz, S. Goericke [et al.] // *Radiology*. – 2007. – Vol. 243. – № 3. – P. 720-726.
74. Jha B. Pearls & Oysters: Hyperdense or pseudohyperdense MCA sign: A Damocles sword? / B. Jha, M. Kothari // *Neurology*. – 2009. – Vol. 72. – № 23. – P. e116-e117.
75. Kobkitsuksakul C. Interobserver agreement between senior radiology resident, neuroradiology fellow, and experienced neuroradiologist in the rating of Alberta Stroke Program Early Computed Tomography Score (ASPECTS) / C. Kobkitsuksakul, O. Tritanon, V. Suraratdecha // *Diagnostic and Interventional Radiology*. – 2018. – URL: http://cms.galenos.com.tr/Uploads/Article_55149/Diagn%20Interv%20Radiol-24-104-En.pdf (date accessed: 22.08.2023).
76. Koo C. K. What Constitutes a True Hyperdense Middle Cerebral Artery Sign? / C. K. Koo, E. Teasdale, K. W. Muir // *Cerebrovascular Diseases*. – 2000. – Vol. 10. – № 6. – P. 419-423.
77. Kumar R. Receiver operating characteristic (ROC) curve for medical researchers / R. Kumar, A. Indrayan // *Indian Pediatrics*. – 2011. – Vol. 48. – № 4. – P. 277-287.
78. Long-Term Prediction of Functional Outcome After Stroke Using the Alberta Stroke Program Early Computed Tomography Score in the Subacute Stage / L. D. Alexander, J. A. Pettersen, J. J. Hopyan [et al.] // *Journal of Stroke and Cerebrovascular Diseases*. – 2012. – Vol. 21. – № 8. – P. 737-744.
79. Mainali S. Detection of Early Ischemic Changes in Noncontrast CT Head Improved with ‘Stroke Windows’ / S. Mainali, M. Wahba, L. Eljovich // *ISRN Neuroscience*. – 2014. – Vol. 2014. – P. 1-4.
80. Matrix metalloproteinase-9 and urokinase plasminogen activator mediate interleukin-1-induced neurotoxicity / P. Thornton, E. Pinteaux, S. M. Allan, N. J.

- Rothwell // *Molecular and Cellular Neuroscience*. – 2008. – Vol. 37. – № 1. – P. 135-142.
81. Minds treating brains: understanding the interpretation of non-contrast CT ASPECTS in acute ischemic stroke / A. T. Wilson, S. Dey, J. W. Evans [et al.] // *Expert Review of Cardiovascular Therapy*. – 2018. – Vol. 16. – № 2. – P. 143-153.
82. Multi-center prediction of hemorrhagic transformation in acute ischemic stroke using permeability imaging features / F. Scalzo, J. R. Alger, X. Hu [et al.] // *Magnetic Resonance Imaging*. – 2013. – Vol. 31. – № 6. – P. 961-969.
83. Multiclass Support Vector Machine-Based Lesion Mapping Predicts Functional Outcome in Ischemic Stroke Patients / N. D. Forkert, T. Verleger, B. Cheng [et al.] // *PLOS ONE*. – 2015. – Vol. 10. – № 6. – P. e0129569.
84. Multiple Sclerosis Identification by 14-Layer Convolutional Neural Network with Batch Normalization, Dropout, and Stochastic Pooling / S.-H. Wang, C. Tang, J. Sun [et al.] // *Frontiers in Neuroscience*. – 2018. – Vol. 12. – P. 818.
85. Navarro-Orozco D. Neuroanatomy, Middle Cerebral Artery / D. Navarro-Orozco, J. C. Sánchez-Manso. – Text: electronic // *StatPearls*. – Treasure Island (FL): StatPearls Publishing, 2023. – URL: <http://www.ncbi.nlm.nih.gov/books/NBK526002/> (date accessed: 22.08.2023).
86. Neutrophil Cerebrovascular Transmigration Triggers Rapid Neurotoxicity through Release of Proteases Associated with Decondensed DNA / C. Allen, P. Thornton, A. Denes [et al.] // *The Journal of Immunology*. – 2012. – Vol. 189. – № 1. – P. 381-392.
87. Nogles T. E. Middle Cerebral Artery Stroke / T. E. Nogles, M. A. Galuska. – Text: electronic // *StatPearls*. – Treasure Island (FL): StatPearls Publishing, 2023. – URL: <http://www.ncbi.nlm.nih.gov/books/NBK556132/> (date accessed: 22.08.2023).
88. Obuchowski N. A. Statistical considerations for testing an AI algorithm used for prescreening lung CT images / N. A. Obuchowski, J. A. Bullen // *Contemporary Clinical Trials Communications*. – 2019. – Vol. 16. – P. 100434.

89. Performance of e-ASPECTS software in comparison to that of stroke physicians on assessing CT scans of acute ischemic stroke patients / C. Herweh, P. A. Ringleb, G. Rauch [et al.] // *International Journal of Stroke*. – 2016. – Vol. 11. – № 4. – P. 438-445.
90. Prakkamakul S. ASPECTS CT in Acute Ischemia: Review of Current Data / S. Prakkamakul, A. J. Yoo // *Topics in Magnetic Resonance Imaging*. – 2017. – Vol. 26. – № 3. – P. 103-112.
91. Predicting Stroke Outcome Using Clinical- versus Imaging-based Scoring System / J. H. Baek, K. Kim, Y.-B. Lee [et al.] // *Journal of Stroke and Cerebrovascular Diseases*. – 2015. – Vol. 24. – № 3. – P. 642-648.
92. Prediction of Hemorrhagic Transformation Severity in Acute Stroke From Source Perfusion MRI / Y. Yu, D. Guo, M. Lou [et al.] // *IEEE Transactions on Biomedical Engineering*. – 2018. – Vol. 65. – № 9. – P. 2058-2065.
93. Prediction of stroke thrombolysis outcome using CT brain machine learning / P. Bentley, J. Ganesalingam, A. L. Carlton Jones [et al.] // *NeuroImage: Clinical*. – 2014. – Vol. 4. – P. 635-640.
94. Prediction of Tissue Outcome and Assessment of Treatment Effect in Acute Ischemic Stroke Using Deep Learning / A. Nielsen, M. B. Hansen, A. Tietze, K. Mouridsen // *Stroke*. – 2018. – Vol. 49. – № 6. – P. 1394-1401.
95. Recommendations on telestroke in Europe / A. M. Alasheev, G. J. Hubert, G. C. Santo [et al.] // *Zhurnal nevrologii i psikiatrii im. S.S. Korsakova*. – 2020. – Vol. 120. – № 3. – P. 33.
96. Reliability and accuracy of individual Alberta Stroke Program Early CT Score regions using a medical and a smartphone reading system in a telestroke network / A. J. Salazar, N. Useche, M. Granja [et al.] // *Journal of Telemedicine and Telecare*. – 2021. – Vol. 27. – № 7. – P. 436-443.
97. Reliability of Visual Assessment of Non-Contrast CT, CT Angiography Source Images and CT Perfusion in Patients with Suspected Ischemic Stroke / T. Van

- Seeters, G. J. Biessels, J. M. Niesten [et al.] // PLoS ONE. – 2013. – Vol. 8. – № 10. – P. e75615.
98. Role of imaging in early diagnosis of acute ischemic stroke: a literature review / M. A. Akbarzadeh, S. Sanaie, M. Kuchaki Rafsanjani, M.-S. Hosseini // The Egyptian Journal of Neurology, Psychiatry and Neurosurgery. – 2021. – Vol. 57. – № 1. – P. 175.
99. Sakai K. Machine learning studies on major brain diseases: 5-year trends of 2014–2018 / K. Sakai, K. Yamada // Japanese Journal of Radiology. – 2019. – Vol. 37. – № 1. – P. 34-72.
100. Samuel's Checkers Player / E. Martin, S. Kaski, F. Zheng [et al.]. – Text: electronic // Encyclopedia of Machine Learning / eds. C. Sammut, G. I. Webb. – Boston, MA: Springer US, 2011. – P. 881-881. – URL: http://link.springer.com/10.1007/978-0-387-30164-8_740 (date accessed: 22.08.2023).
101. Schröder J. A Critical Review of Alberta Stroke Program Early CT Score for Evaluation of Acute Stroke Imaging / J. Schröder, G. Thomalla. // Frontiers in Neurology. – 2017. – Vol. 7. – URL: <http://journal.frontiersin.org/article/10.3389/fneur.2016.00245/full> (date accessed: 22.08.2023).
102. Shen D. Deep Learning in Medical Image Analysis / D. Shen, G. Wu, H.-I. Suk // Annual Review of Biomedical Engineering. – 2017. – Vol. 19. – № 1. – P. 221-248.
103. Streamlining door to recanalization processes in endovascular stroke therapy / A. Aghaebrahim, C. Streib, S. Rangaraju [et al.] // Journal of NeuroInterventional Surgery. – 2017. – Vol. 9. – № 4. – P. 340-345.
104. Stroke care: initial data from a county-based bypass protocol for patients with acute stroke / S. F. Zaidi, J. Shawver, A. Espinosa Morales [et al.] // Journal of Neurointerventional Surgery. – 2017. – Vol. 9. – № 7. – P. 631-635.

105. Sulcal Effacement with Preserved Gray–White Junction: A Sign of Reversible Ischemia / D. C. Haussen, A. Lima, M. Frankel [et al.] // *Stroke*. – 2015. – Vol. 46. – № 6. – P. 1704-1706.
106. The Alberta Stroke Program Early CT Score in clinical practice: what have we learned? / V. Puetz, I. Dzialowski, M. D. Hill, A. M. Demchuk // *International Journal of Stroke: Official Journal of the International Stroke Society*. – 2009. – Vol. 4. – № 5. – P. 354-364.
107. The ASPECTS template is weighted in favor of the striatocapsular region / T. G. Phan, G. A. Donnan, M. Koga [et al.] // *NeuroImage*. – 2006. – Vol. 31. – № 2. – P. 477-481.
108. The validity and reliability of signs of early infarction on CT in acute ischaemic stroke / D. W. J. Dippel, M. Du Ry Van Beest Holle, F. Van Kooten, P. J. Koudstaal // *Neuroradiology*. – 2000. – Vol. 42. – № 9. – P. 629-633.
109. Tilt-Corrected Region Boundaries May Enhance the Alberta Stroke Program Early Computed Tomography Score for Less Experienced Raters / C. J. Culbertson, S. Christensen, M. Mlynash [et al.] // *Journal of Stroke and Cerebrovascular Diseases*. – 2020. – Vol. 29. – № 7. – P. 104820.
110. Time to Treatment with Endovascular Thrombectomy and Outcomes from Ischemic Stroke: A Meta-analysis / J. L. Saver, M. Goyal, A. Van Der Lugt [et al.] // *JAMA*. – 2016. – Vol. 316. – № 12. – P. 1279.
111. Use of the Alberta Stroke Program Early CT Score (ASPECTS) for assessing CT scans in patients with acute stroke / J. H. Pexman, P. A. Barber, M. D. Hill [et al.] // *AJNR. American journal of neuroradiology*. – 2001. – Vol. 22. – № 8. – P. 1534-1542.
112. Using the baseline CT scan to select acute stroke patients for IV-IA therapy / M. D. Hill, A. M. Demchuk, T. A. Tomsick [et al.] // *AJNR. American journal of neuroradiology*. – 2006. – Vol. 27. – № 8. – P. 1612-1616.
113. Validation of automated Alberta Stroke Program Early CT Score (ASPECTS) software for detection of early ischemic changes on non-contrast brain CT scans /

- L. Wolff, O. A. Berkhemer, A. C. G. M. Van Es [et al.] // *Neuroradiology*. – 2021. – Vol. 63. – № 4. – P. 491-498.
114. Validation of Computed Tomographic Middle Cerebral Artery ‘Dot’ Sign: An Angiographic Correlation Study / M. C. Leary, C. S. Kidwell, J. P. Villablanca [et al.] // *Stroke*. – 2003. – Vol. 34. – № 11. – P. 2636-2640.
115. Validity and reliability of a quantitative computed tomography score in predicting outcome of hyperacute stroke before thrombolytic therapy / P. A. Barber, A. M. Demchuk, J. Zhang, A. M. Buchan // *The Lancet*. – 2000. – Vol. 355. – № 9216. – P. 1670-1674.
116. Wardlaw J. M. Early Signs of Brain Infarction at CT: Observer Reliability and Outcome after Thrombolytic Treatment—Systematic Review / J. M. Wardlaw, O. Mielke // *Radiology*. – 2005. – Vol. 235. – № 2. – P. 444-453.
117. Wey H.-Y. A review of current imaging methods used in stroke research / H.-Y. Wey, V. R. Desai, T. Q. Duong // *Neurological Research*. – 2013. – Vol. 35. – № 10. – P. 1092-1102.
118. Why are acute ischemic stroke patients not receiving IV tPA? Results from a national registry / S. R. Messé, P. Khatri, M. J. Reeves [et al.] // *Neurology*. – 2016. – Vol. 87. – № 15. – P. 1565-1574.