

# Качество алгоритмов искусственного интеллекта для выявления признаков рассеянного склероза на магнитно-резонансных томограммах (систематический обзор)

Г.Н. Черняева<sup>1</sup>, С.П. Морозов<sup>1</sup>, А.В. Владимирский<sup>1,2</sup>

<sup>1</sup>ГБУЗ г. Москвы «Научно-практический клинический центр диагностики и телемедицинских технологий Департамента здравоохранения Москвы», Москва, Россия;

<sup>2</sup>ФГАОУ ВО «Первый Московский государственный медицинский университет имени И.М. Сеченова» (Сеченовский Университет), Москва, Россия

Выполнен систематический обзор литературы с целью обобщения данных о точности и результативности применения алгоритмов искусственного интеллекта для выявления рассеянного склероза по результатам магнитно-резонансной томографии. В обзор включены 39 статей, авторами которых предложено большое количество соответствующих алгоритмов и математических моделей. Однако оценка качества таких разработок ограничена ретроспективным тестированием на повторяющихся наборах данных. Практически полностью отсутствуют результаты клинической апробации, нет проспективных независимых научных исследований точности и применимости. Довольно высокие уровни основных метрик (коэффициенты сходства, чувствительность, специфичность — 75–85%) нивелируются методическими ошибками при формировании исходных наборов данных, отсутствием валидации на независимых данных. В силу малых объёмов выборок и методических дефектов оценки точности результаты подавляющего большинства статей не отвечают критериям доказательности. В наиболее качественных, с методической точки зрения, исследованиях достигнута чувствительность алгоритмов 51,6–77,0%, значение коэффициента Дайса–Сёрнсена — 53,5–56,0%. Значение невысоки, но они свидетельствуют о потенциальной реализуемости задачи автоматизированного выявления признаков рассеянного склероза на магнитно-резонансных томограммах. Для дальнейшего развития автоматизированного анализа требуется разработка клинических сценариев применения, формирование методологии тестирования, проведение проспективных клинических апробаций.

**Ключевые слова:** рассеянный склероз; искусственный интеллект; машинное обучение; магнитно-резонансная томография

**Благодарность.** Авторы выражают благодарность к.м.н. Кремневой Елене Игоревне, старшему научному сотруднику ФГБНУ «Научный центр неврологии», за консультации при подготовке статьи.

**Источник финансирования.** Авторы заявляют об отсутствии внешних источников финансирования при проведении исследования.

**Конфликт интересов.** Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

**Адрес для корреспонденции:** 127051, Москва, ул. Петровка, д. 24, стр. 1. ГБУЗ г. Москвы «Научно-практический клинический центр диагностики и телемедицинских технологий Департамента здравоохранения Москвы». E-mail: a.vladimirsky@nrcmr.ru. Владимирский А.В.

**Для цитирования:** Черняева Г.Н., Морозов С.П., Владимирский А.В. Качество алгоритмов искусственного интеллекта для выявления признаков рассеянного склероза на магнитно-резонансных томограммах (систематический обзор). *Анналы клинической и экспериментальной неврологии*. 2021; 15(4): 54–65.

DOI: <https://doi.org/10.54101/ACEN.2021.4.6>

Поступила 07.05.2020 / Принята в печать 16.02.2021

## The quality of artificial intelligence algorithms for identifying manifestations of multiple sclerosis on magnetic resonance imaging (systematic review)

Galina N. Chernyaeva<sup>1</sup>, Sergey P. Morozov<sup>1</sup>, Anton V. Vladzimirsky<sup>1,2</sup>

<sup>1</sup>Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies, Moscow, Russia;

<sup>2</sup>I.M. Sechenov First Moscow State Medical University (Sechenov University), Moscow, Russia

A systematic review was undertaken to summarize the data regarding accuracy and effectiveness of artificial intelligence algorithms for identifying MRI manifestations of multiple sclerosis. The review included 39 papers, whose authors put forth a multitude of corresponding algorithms and mathematical models. However, quality assessment of these developments was limited by retrospective testing on repeat data sets. Clinical test results were almost entirely absent, and there were

*no prospective independent studies of accuracy and applicability. The relatively high values obtained for the main measures (similarity, sensitivity and specificity coefficients, which were 75–85%) were offset by the methodological errors when creating the baseline data sets, and lack of validation using independent data. Due to small sample sizes and methodological errors when measuring the result accuracy, most of the studies did not meet the criteria for evidence-based research. Studies with the highest methodological quality had algorithms that achieved a sensitivity of 51.6–77.0%, with a S resen–Dice coefficient of 53.5–56.0%. These numbers are not high, but they indicate that automatic identification of multiple sclerosis manifestations on magnetic resonance imaging may be achievable. Further development of computer-aided analysis requires the creation of clinical use scenarios and testing methodology, and prospective clinical testing.*

**Keywords:** multiple sclerosis; artificial intelligence; machine learning; magnetic resonance imaging

**Acknowledgements.** The authors are grateful to Elena I. Kremneva, Candidate of Medical Sciences and Senior Researcher at the Neurology Research Centre, for her advice during the preparation of this article.

**Source of funding.** This study was not supported by any external sources of funding.

**Conflict of interest.** The authors declare no apparent or potential conflicts of interest related to the publication of this article.

**For correspondence:** 127051, Moscow, Petrovka Str., 24, bld. 1. Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies of the Moscow Health Care Department. E-mail: a.vladimirsky@npcmr.ru. Vladzymyrsky A.V.

**For citation:** Chernyaeva G.N., Morozov S.P., Vladzymyrsky A.V. [The quality of artificial intelligence algorithms for identifying manifestations of multiple sclerosis on magnetic resonance imaging (systematic review)]. *Annals of clinical and experimental neurology*. 2021; 15(4): 55–65. (In Russ.)

DOI: <https://doi.org/10.54101/ACEN.2021.4.6>

Received 07.05.2020 / Accepted 16.02.2021

## Введение

Среди хронической патологии центральной нервной системы особое место занимает рассеянный склероз (РС) — неизлечимое, аутоиммунное и нейродегенеративное заболевание, являющееся одной из основных причин стойкой утраты трудоспособности у лиц молодого возраста. В последнее время отмечается глобальный рост распространённости РС [1–3]. Далеко не последнюю роль в увеличении выявляемости РС играет интенсивное развитие диагностической нейровизуализации. Во многих странах Европы, Юго-Восточной Азии, Северной Америки наблюдается прогрессивный рост оснащённости систем здравоохранения компьютерными и магнитно-резонансными томографами с параллельным увеличением числа исследований. Однако количество диагностических процедур не всегда коррелирует с качеством диагностики. Ряд авторов полагает, что увеличение распространённости РС ошибочно и обусловлено гипердиагностикой этого состояния в 10–31% случаев [4–6]. Опубликованные протипированными авторами данные свидетельствуют о низком качестве диагностики РС на первичном уровне здравоохранения, причём эта ситуация носит глобальный характер. Гипердиагностика приводит к бессмысленному, высокочувствительному лечению, психологическим расстройствам, стигматизации. А гиподиагностика, наоборот, задерживает начало специфической терапии, приближает инвалидизацию, ухудшает прогноз. В связи с этим поиск путей повышения качества и точности диагностики РС на первичном уровне медицинской помощи является актуальной задачей.

С учётом уровня развития цифровых технологий представляется потенциально возможным применять алгоритмы искусственного интеллекта (ИИ) для автоматизированного анализа медицинской информации и поддержки принятия решений врачами-радиологами [7].

Предварительный анализ литературы позволил установить, что разработки в сфере ИИ для диагностики РС ведутся с 1990-х гг. [8–14]: предлагались различные ма-

тематические модели и методы анализа данных, однако валидация этих разработок практически полностью отсутствовала, а стандартные метрики точности не публиковались. Период исследования проблемы с 1990-х до 2010-х гг. внёс вклад в развитие математики и компьютерных наук, но не медицины.

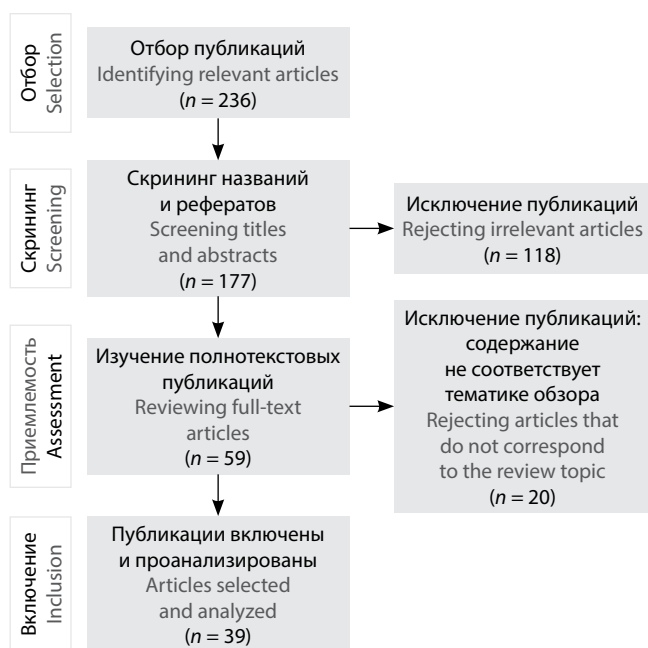
Позднее для скрининга и прогнозирования течения болезни предлагались:

1. Алгоритмы обработки естественного языка для анализа медицинской документации (истории болезни, результаты лабораторных исследований, антропологические и демографические данные и пр.) [15–18]. Во всех указанных работах авторы заявляли о хороших результатах, но независимого тестирования, масштабирования и клинической апробации этих разработок не проводилось. Иногда такие разработки дополнялись средствами анализа магнитно-резонансных изображений, что обеспечивало чувствительность и специфичность метода на уровне 71% и 68% соответственно [19].
2. Алгоритмы анализа результатов электроэнцефалографии [20–22]. С точки зрения автоматизации, это направление является бесперспективным, т.к. электроэнцефалография не позволяет провести диагностику РС в соответствии с критериями McDonald. Лидирующим методом, безусловно, является магнитно-резонансная томография (МРТ).
3. Алгоритмы анализа результатов МРТ посредством различных математических методов, в основном нейронных сетей. Именно это направление представляется наиболее перспективным и требующим углубленного изучения.

**Цель исследования** — систематизировать данные о точности и результативности применения алгоритмов ИИ для диагностики РС по результатам МРТ.

## Материалы и методы

Выполнен систематический обзор с описанием по методологии «PRISMA» [23].



### Стратегия отбора публикаций для систематического обзора.

Article selection process for the systematic review.

Поиск публикаций проводился на двух языках (русском и английском) с использованием ресурсов Научной электронной библиотеки РФ ([www.elibrary.ru](http://www.elibrary.ru)), Национальной медицинской библиотеки США ([www.pubmed.org](http://www.pubmed.org)), коллекции препринтов Корнеллского университета, Итака, США ([www.arxiv.org](http://www.arxiv.org)).

Поисковые запросы на английском языке включали термины (медицинские предметные заголовки (Medical Subject Headings — MeSH): «multiple sclerosis», «artificial intelligence», «machine learning». Поисковые запросы на русском языке включали термины: «рассеянный склероз», «искусственный интеллект», «автоматизированный анализ», «машинное обучение».

Стратегия поиска представлена на рисунке. Обзор актуален по состоянию на 01.04.2020.

Критерии включения публикации:

- 1) соответствие теме систематического обзора;
- 2) оригинальное исследование (дизайн диагностического исследования);
- 3) статья в рецензируемом журнале, серии рецензируемых сборников, препринт;
- 4) в обучающий дата-сет включены результаты обследования пациентов исключительно с РС и лиц без патологии головного мозга;
- 5) алгоритм проводит анализ только результатов МРТ головного мозга (сегментация структур, выявление очагов);
- 6) приведены объективные данные о валидации, диагностической точности, иных метриках эффективности алгоритмов;
- 7) для объективизации диагностической точности использованы стандартные метрики (чувствительность, специфичность, площадь под характеристической кривой, коэффициент Дайса—Сёренсена).

Далее из каждой публикации были извлечены необходимые данные:

- имя автора, год публикации;
- цель, дизайн исследования;
- данные о наборе данных (дата-сете), методике проведения МРТ;
- результат оценки точности и эффективности;
- общие результаты.

Перечисленные сведения были систематизированы и проанализированы.

### Результаты и обсуждение

В обзор были включены 39 статей и препринтов. Подавляющая их часть представляет собой детальное описание математических аспектов разработки алгоритмов (моделей) для автоматизированного выявления признаков РС на МРТ; при этом процесс и результаты оценки точности приводятся крайне лаконично. Нами не выявлено ни одного исследования, в котором алгоритм применялся бы в условиях клинической апробации или для проспективного анализа изображений. Все статьи, подходящие по критериям включения, имели ретроспективный дизайн. В публикациях отсутствовали данные о клинических результатах, поэтому мы ограничились систематизацией вопросов, связанных с наборами данных (дата-сетями), методикой и результатами ретроспективной оценки диагностической точности. Обобщённая информация приведена в таблице 1.

#### Наборы данных (дата-сеты)

Для разработки алгоритмов выявления РС по результатам МРТ группы исследователей использовали собственные и публичные наборы данных. Из числа публичных дата-сетов наиболее часто использовался «MICCAI 2008» — с ним работали 5 групп разработчиков, «ISBI 2015» — 3, «MS-100», «ISBI-61», «Bernese MS cohort» — по 1.

В выборки включались данные пациентов обоих полов старше 18 лет. Информация о демографических данных лиц, чьи исследования были включены в дата-сеты, представлена в 18 (46,2%) статьях, сведения о типе и клинических особенностях основного заболевания — в 13 (33,3%). При этом ремиттирующий и первично-прогрессирующий варианты течения встречались примерно в равном количестве статей.

В результате анализа публикаций установлена характерная ситуация: большинство разработчиков в наборы данных включали только целевые случаи с РС. Можно сказать, что их алгоритмы «оставались в неведении» о существовании изображений без признаков РС. Такой подход является грубейшей методической ошибкой. Лишь 8 (20,5%) разработчиков использовали наборы данных, включающие результаты МРТ здоровых лиц или исследования с сосудистыми очагами [24–30]. В наборы данных были включены следующие режимы сканирования: T1-взвешенные изображения (T1-ВИ), T2-ВИ, T2-ВИ с подавлением сигнала от свободной жидкости (англ. Fluid Attenuation Inversion Recovery (FLAIR)), изображения, взвешенные по протонной плотности (англ. Proton Density Weighted (PDW)). Большинство разработчиков использовали 3 вида изображений (T1-ВИ, T2-ВИ, FLAIR) — 12 (30,8%), только FLAIR — 6 (15,4%), все 4 вида изображений — 7 (18,0%). Сочетание

**Обобщение данных об алгоритмах ИИ для выявления признаков РС на МРТ**  
 Summarizing data on AI algorithms to detect signs of multiple sclerosis in MRI

Автор, год публикации Author, year of publication	Набор данных Data set			коэффициент Дайса-Сёрнсена Sørensen-Dice coefficient	Метрики точности Precision metrics			Примечание Comments
	источник source	число пациентов с РС number of patients with MS	число пациентов без РС number of patients without MS		чувствительность sensitivity	специфичность specificity	общая точность general accuracy	
Salem M. <i>et al.</i> , 2017 [43]	Собственный Own	36	-	0,77 ± 0,23, 0,56 ± 0,23	74,30 ± 28,70	11,86 ± 18,40		
Brosch T. <i>et al.</i> , 2016 [37]	«МССАИ 2008», «ISBI 2015» Собственный Own	377	-		51,6		51,3*	Тестирование на новых данных Tested with updated data
Gabr R.E. <i>et al.</i> , 2019 [36]	Собственный Own	1008	-	0,82 (95% ДИ 0,61–1,0)	76,0–79,0			Мультицентровое исследование (данные из 68 организаций). Максимальная чувствительность ≥ 91% достигается для очагов объёмом не менее 70 мм³ Multicenter study (data provided by 68 organizations) Highest responsiveness value ≥ 91% was observed for focal areas of at least 70 mm³
Wang S.H. <i>et al.</i> , 2018 [24]	То же Ibid	38	26	98,77 ± 0,35	98,76 ± 0,58		98,77 ± 0,39 98,75 ± 0,58**	
Zurita M. <i>et al.</i> , 2018 [25]	- -	104	46				89,0 ± 2,0, 63,0 ± 5,0	Сравнение пациентов с разной степенью неврологического дефицита Comparison of patients with varied neurologic deficit
Yoo Y. <i>et al.</i> , 2017 [26]	- -	55	44		87,3 ± 12,9	88,6 ± 12,5	87,9 ± 8,4 88,0 ± 8,5***	
Valverde S. <i>et al.</i> , 2017 [38]	«МССАИ 2008» Собственный Own	105	-	53,5; 56,0	55,5–68,7; 77,0; 68,2		70,3; 73,0**	Тестирование на новых данных Tested with updated data

Продолжение на странице 58.

Автор, год публикации Author, year of publication	Набор данных Data set			Метрики точности Precision metrics			Примечание Comments
	источник source	число пациентов с РС number of patients with MS	число пациентов без РС number of patients without MS	чувствительность sensitivity	специфичность specificity	общая точность general accuracy	
Goldberg-Zimmering D. <i>et al.</i> , 1998 [41]	Собственный Own	14	-	0,87	0,96		
Boudraa A.O. <i>et al.</i> , 2000 [44]	Собственный Own	10	-			0,65–0,87****	
Geremia E. <i>et al.</i> , 2010 [45]	«МИССАИ 2008»	20	-	0,23–0,49		0,28–0,78**	
Yamamoto D. <i>et al.</i> , 2010 [40]	Собственный Own	3	-	81,5			
Shiee N. <i>et al.</i> , 2010 [46]	То же Ibid	10	-	0,633	0,712		
García-Lorenzo D. <i>et al.</i> , 2009 [47]	- "-	10	-	0,63			Мультицентровое исследование (данные из 3 организаций) Multicenter study (data provided by 3 organizations)
Cabezas M. <i>et al.</i> , 2014 [35]	- "-	45	-	0,44–0,56			
Crimi A. <i>et al.</i> , 2014 [31]	- "-	25	-			0,80 ± 0,13, 0,90 ± 0,05 (коэффициент детерминации / R <sup>2</sup> coefficient) / provided by 5 organizations)	Мультицентровое исследование (данные из 5 организаций) Multicenter study (data provided by 5 organizations)
Theodorakis P. <i>et al.</i> , 2009 [30]	- "-	11	18			88,46%	В дата-сет включены исследования с сосудистыми очагами Dataset includes data on focal areas of high vascular density
Khayati R. <i>et al.</i> , 2008 [49]	- "-	20	-	0,74 ± 0,05		0,75 ± 0,03****	
Weiss N. <i>et al.</i> , 2013 [50]	«МИССАИ 2008»	20	-	63,0–71,0	33,0	37,0**	
Gao J. <i>et al.</i> , 2014 [51]	«МИССАИ 2008»	20	-	0,55–0,59	0,98–0,99		

Продолжение на странице 59.

Автор, год публикации Author, year of publication	Набор данных Data set			Метрики точности Precision metrics			Примечание Comments		
	источник source	число пациентов с РС number of patients with MS	число пациентов без РС number of patients without MS	коэффициент Дайса-Сёрсенена Sørensen-Dice coefficient	чувствительность sensitivity	специфичность specificity		общая точность general accuracy	иное other
Nakamura K. <i>et al.</i> , 2009 [27]	Собственный Own	3	3					0,836****	
Bijar A. <i>et al.</i> , 2013 [52]	То же Ibid	20	-					0,7132-0,8262****	
Cerasa A. <i>et al.</i> , 2012 [53]	- -	11	-	19,7 ± 12,01					
Hackmack K. <i>et al.</i> , 2012 [28]	- -	41	26		92,68	96,15	94,42		
Aumerich F.X. <i>et al.</i> , 2011 [54]	- -	4	-		0,760				
Bendfeldt K. <i>et al.</i> , 2012 [42]	- -	50	-		76,9-82,3	76,9-88,2			Сравнение пациентов, получающих и не получающих медикаментозную терапию Comparison of patients who received drug therapy versus watch-and-wait patients
Kuwazuru J. <i>et al.</i> , 2012 [56]	- -	3	-		75,0-86,3			0,17-0,49****	
Jain S. <i>et al.</i> , 2016 [34]	- -	22	-	0,60-0,89	0,42-0,57			0,25-0,48*	В дата-сет включены исследования, выполненные на 3 разных томографах. Данный алгоритм входит в программное обеспечение «Icometrix», сертифицированное в США как медицинское изделие The data set included images obtained from 3 different CT scanners. This algorithm is used by Icometrix software, certified as a medical device in the USA
Kanber B. <i>et al.</i> , 2019 [56]	- -	124	-		77,7 (95% ДИ) 77,2-78,3	77,8 (95% ДИ) 77,3-78,2	86,5	0,890****	

Продолжение на странице 60.

Автор, год публикации Author, year of publication	Набор данных Data set			Метрики точности Precision metrics			Примечание Comments
	источник source	число пациентов с РС with MS	число пациентов без РС number of patients without MS	чувствительность sensitivity	специфичность specificity	общая точность general accuracy	
Wei W, <i>et al.</i> , 2018 [8]	- -	18	10	0,83			
Freire P.G.L. <i>et al.</i> , 2018 [57]	- -	5	-	0,9860 ± 0,0012			
Feng Y. <i>et al.</i> , 2018 [58]	«ISBI 2015»	14	-	0,684	0,648		0,832**
Kazanci E. <i>et al.</i> , 2018 [59]	Собственный Own	9	-	57,5 ± 12,4	59,7 ± 14,6		
Roy S. <i>et al.</i> , 2018 [60]	«ISBI 2015» собственный Own	100	-	0,5639			0,6040** Тестирование на новых данных Tested with updated data
Sacca V. <i>et al.</i> , 2019 [61]	Собственный Own	18	19	46,6–100,0	46,6–100,0	56,5–85,7	
Zhang H. <i>et al.</i> , 2020 [62]	То же Ibid	43	-	66,011			
Gheshlaghi S.H. <i>et al.</i> , 2019 [63]	- -	35	35			0,99	
McKinley R. <i>et al.</i> , 2019 [33]	«Bernese MS cohort databank»	50	-		0,38–1,00	0,52–0,94	0,77–0,99*** Мультицентровое исследование (данные из 2 организаций). Тестирование на новых данных Multicenter study (data provided by 2 organizations). Tested with updated data
Aslani S. <i>et al.</i> , 2019 [32]	«ISBI 2015» Собственный Own	56	-	0,6655–0,7067			0,6844– 0,8032**
Salem M. <i>et al.</i> , 2020 [48]	Собственный Own	60	-	0,83	83,09		9,36*

Примечание. \*\*Частота ложноположительных случаев; \*\*предсказательная ценность положительного результата; \*\*\*площадь под характеристической кривой; \*\*\*\*коэффициент подобия.  
Note. \*Frequency of false-positive outcomes; \*\* prognostic value of positive findings; \*\*\*area under H and D curve; \*\*\*\*similarity index.

FLAIR только с T1-ВИ или только с T2-ВИ сочли оптимальным 6 (15,4%) разработчиков. T1-ВИ и T2-ВИ применили 7 (18,0%). Остальные руководствовались иными сочетаниями видов изображений (однократные случаи). Одна группа авторов ограничилась только T1-ВИ.

Примерно в равном количестве статей использованы результаты МРТ с толщиной срезов 1 или 3 мм.

В 15 (38,5%) работ исследования проводились на томографах с напряжённостью магнитного поля 3 Т, в 10 (25,7%) — 1,5 Т (большинство этих работы выполнены до 2013 г.). 5 (12,8%) исследователей в дата-сетях смешали исследования с аппаратов 1,5 и 3 Т. В 9 (23,0%) статьях данные о характеристиках томографов не представлены.

Данные об аппаратах, проекциях, иных деталях выполнения МРТ авторами статей приведены хаотично, что не позволяет выполнить обобщение.

### Диагностическая точность

Подавляющее большинство авторов — 34 (87,2%) — использовали подход «leave-one-out», суть которого состоит в разделении имеющегося набора данных на выборку для обучения и выборку для оценки точности. При этом общий объём наборов данных колебался от 3 до 150 клинических случаев (в среднем 38 случаев при значениях медианы 23,5 и моды 20); 3 (7,7%) разработчика использовали более 100 случаев, 9 (23,1%) — 10 и менее.

Только 5 групп разработчиков использовали дата-сет, сформированные из данных нескольких медицинских организаций: А. Crimi и соавт. — 25 случаев из 5 клиник [31], S. Aslani и соавт. — 56 случаев из собственного и из публичного дата-сета [32], R. McKinley и соавт. — 50 случаев из 2 медицинских организаций [33], S. Jain и соавт. — 22 случая с 3 разных томографов [34], M. Cabezas и соавт. — 45 случаев из 3 клиник [35].

Валидация алгоритмов на независимых данных проведена 5 (12,8%) разработчиками; причём для обучения алгоритмов использованы наиболее значительные объёмы данных (100 и более случаев). Для разработки и валидации R.E. Gabr и соавт. использовали набор данных из 1008 случаев, собранных из 68 медицинских организаций [36]. Отметим, что это самый объёмный дата-сет, использованный для обучения и валидации алгоритмов ИИ, направленных на выявление РС. R. McKinley и соавт. применяли два различных дата-сета из двух клиник: один — для обучения, второй — для тестирования [33]. Три группы исследователей использовали собственные наборы (37, 60 и 77 случаев) для обучения, а релевантные дата-сеты, находящиеся в открытом доступе (MISCAI 2008, ISBI 2015) — для тестирования точности [32, 37, 38].

Таким образом, доказательность показателей диагностической точности, заявляемых 87,2% разработчиков, крайне сомнительна. Вызывает недоумение использование для обучения и валидации дата-сетов из 3, 4, 5, 9, 10, 11, 14 случаев, а также исключительное обучение на результатах исследований с признаками РС. В таких ситуациях происходит фактически экспериментальная разработка алгоритмов, совершенно «оторванная» от реальных клинических условий. Минимальное количество разработчиков (5 (12,8%)) осознали необходимость мультицентрового подхода,

т.е. формирования дата-сетов из данных нескольких медицинских организаций, тестирования разработок на независимых (новых для алгоритма) изображениях.

Оценку диагностической точности авторы статей проводили путем вычисления следующих показателей:

- коэффициенты сходства (Дайса–Сёренсена, индекс сходства);
- чувствительность и специфичность;
- общая точность;
- предсказательная ценность положительного или отрицательного результата;
- частота ложноположительных случаев (false positive rate);
- площадь под характеристической кривой;
- коэффициент детерминации.

Вместе с тем ни в одной статье не был соблюден дизайн диагностического исследования [39]. Вычисления метрик авторы проводят бессистемно, дискретно. Например, приводятся данные о чувствительности, но отсутствует специфичность. Иногда есть попытки ввести собственные коэффициенты, видоизменить стандартные формулы. Такой подход нарушает принципы выполнения биомедицинских исследований. Для медицинской аудитории доказательность соответствующих публикаций минимальна. Большинство авторов ( $n = 20$ ; 51,3%) использовали коэффициент Дайса–Сёренсена. Еще в 6 работах использован индекс сходства. Корректно привели данные о чувствительности и специфичности только 9 (23,1%) групп исследователей. Площадь под характеристической кривой использовали только 2 группы авторов. В целом для алгоритмов выявления РС на МРТ значения коэффициентов сходства колеблются в очень широком диапазоне (0,44–0,99); аналогичная ситуация для чувствительности (0,23–0,99), специфичности (0,12–0,99), общей точности (0,58–0,99).

С учётом методических дефектов и нерепрезентативности большинства дата-сетов (как было показано выше) проводить математическую обработку и метаанализ этих данных не имеет смысла.

Таким образом, качество включённых в обзор статей с позиций дизайна является низким. Представление показателей диагностической точности не соответствует принципам доказательной медицины. Довольно высокие уровни основных метрик (коэффициенты сходства, чувствительность, специфичность) нивелируются методическими ошибками при формировании исходных дата-сетов, отсутствием валидации на независимых данных. В силу малых объёмов выборки и методических дефектов оценки точности результаты подавляющего большинства статей не отвечают критериям доказательности.

Одна из наиболее методически качественных работ — статья S. Valverde и соавт. [38]. Авторы предложили каскад свёрточных нейронных сетей для сегментации белого вещества и выявления типичных для РС поражений головного мозга. Для обучения модели и ее тестирования «leave-one-out» использован «экспериментальный» набор данных (публично доступный дата-сет «MISCAI 2008»;  $n = 45$ ), а для независимой валидации — два собственных «клинических» набора ( $n = 35$ ;  $n = 25$ ), которые не были «известны» модели до тестирования. Чувствительность алгоритма на «экспериментальном» дата-сете составила 55,5–68,7%,



на «клинических» — 77,0 и 68,2% соответственно. Значения коэффициента Дайса—Сёрнсена для «клинических» наборов данных составили 53,5 и 56,0%, превосходя результаты для тестирования на «экспериментальном» дата-сете. Также авторы сравнили собственные результаты и данные литературы [38]. Полученные сведения наглядно продемонстрировали не только превосходство модели над аналогами, но и хорошую воспроизводимость результатов работы алгоритма. С клинической точки зрения, достигнутые значения метрик точности низки и вряд ли позволяют говорить даже об апробации в условиях практического здравоохранения; явно требуется дополнительное обучение модели. Но авторы придерживаются очень правильного методического подхода как к разработке алгоритмов, так и к их тестированию. Аналогичный подход применили Т. Brosch и соавт., используя собственный дата-сет для обучения, а два публичных («MISCAI 2008» и «ISBI 2015») — для тестирования своего алгоритма на новых данных [17]. Всего авторы использовали 377 случаев, достигнув чувствительность 51,6% и предсказательную ценность отрицательного результата 51,3%.

Р.Е. Gabr и соавт. реализовали модель на основе свёрточной нейронной сети с полностью связанными слоями [36]. В мультицентровом исследовании авторы использовали дата-сет из 1008 случаев, собранных из 68 медицинских организаций. Разработчики смогли получить значение коэффициента Дайса—Сёрнсена 0,82 (95% ДИ 0,61–1,0). Особый интерес, с точки зрения методологии разработки и применения алгоритмов ИИ для выявления РС, представляет следующий факт. Авторы смогли убедительно доказать, что наибольшая чувствительность алгоритма (91% и более) достигается для очагов объемом 70 мм<sup>3</sup>.

Примечательно, что в описанных работах получены довольно скромные значения метрик диагностической точности. При этом в исследовании по принципу «leave-one-out» чувствительность достигает 81,5% (обучение и тестирование на 3 случаях) [40], 87,0% (на 14 случаях) [41], 92,68% (на 67 случаях) [28], 98,77 (на 64 случаях) [24]. Аналогичная ситуация и с коэффициентами сходства. Без проверки точности на новых данных, без независимой валидации результатов у разработчиков создаётся иллюзия высокого качества алгоритмов, возможности «замены рентгенологов» с их помощью. Особо удивляет такая позиция при обучении и одновременной валидации алгоритмов на 3–9 изображениях. Принцип «leave-one-out» подходит для предварительной оценки точности на начальных этапах разработки, но по мере эволюции алгоритма для его тестирования надо обязательно применять новые данные, полученные на разных диагностических устройствах и в разных медицинских организациях.

Особо надо отметить две статьи, в которых сравнивались работы алгоритмов у разных групп пациентов: получавших

или не получавших медикаментозную терапию препаратами, изменяющими течение РС [42]; имеющих различную степень неврологического дефицита по расширенной шкале оценки степени инвалидизаций Куртцке (англ. Expanded Disability Status Scale) [25]. В обоих исследованиях достигнуты средние уровни точности, свидетельствующие о необходимости дальнейшей работы над обучением алгоритмов. Однако выделение авторами разных клинических групп является правильным и перспективным подходом.

## Выводы

1. Над задачей выявления РС путём автоматизированного анализа МРТ работает значительное число исследователей; предложено большое количество математических методов.
2. Предварительные результаты обнадеживают. Значения ключевых метрик диагностической точности в среднем достигают 75–85%. Это позволяет говорить о потенциальной реализуемости задачи автоматизированного выявления признаков РС на МРТ, а также о необходимости углублённых исследований.
3. Вместе с тем на момент подготовки обзора отсутствуют результаты клинической апробации алгоритмов ИИ для выявления РС, нет проспективных научных исследований точности и применимости. Независимое тестирование алгоритмов осуществляется только в единичных исследованиях; тогда как валидация на новых данных должна быть рутинной частью процесса разработки.
4. Для дальнейшего качественного развития алгоритмов ИИ для выявления и мониторинга РС требуется разработка клинических сценариев их применения, формирование методологии объективного тестирования (на принципах доказательной медицины, со стандартизацией критериев), проведение проспективных клинических апробаций.
5. Требуется обоснование стандартных требований к наборам данных, используемых для обучения алгоритмов, включая вопросы стандартизации, методологии подготовки (разметки) изображений, юридические аспекты.
6. Применение алгоритмов ИИ в контексте проблемы РС может рассматриваться по следующим направлениям:
  - поддержка принятия решений по результатам первого МРТ (выявление характерных признаков РС, определение типа патологического процесса, прогнозирование);
  - дифференциальная диагностика в сложных случаях (возможно с использованием клинических данных);
  - подбор и оценка результативности терапии, выявление ранних признаков осложнений;
  - контроль динамики патологического процесса.

Перечисленные направления должны быть разделены на отдельные проблемы, что, в том числе, будет являться предметом наших дальнейших исследований.

## Список источников

1. Абдурахманова Р.Ф., Иззатов Х.Н., Халибаева Г.Р., Шарипова Б.А., Кахарова М.Х. Рассеянный склероз: этиология, патогенез и клиника (часть I). *Вестник последипломного образования в сфере здравоохранения*. 2016;(3):68–74.
2. Попова Е.В., Бойко А.Н., Барабанова М.А. и соавт. Первично-прогрессирующий рассеянный склероз: современное состояние проблемы своевременной постановки диагноза. *Журнал неврологии и психиатрии им. С.С. Корсакова*. 2017;117(10–12):35–40. DOI: 10.17116/jnevro201711710235-40.

## References

1. Abdurakhmanova R.F., Izzatov Kh.N., Khadibaeva G.R. et al. Multiple sclerosis: etiology, pathogenesis and clinics (part I). *Vestnik poslediplomnogo obrazovaniya v sfere zdravookhraneniya*. 2016;(3):68–74. (In Russ).
2. Popova E.V., Boyko A.N., Barabanova M.A. et al. Primary progressive multiple sclerosis: current issues of timely diagnosis. *Zhurnal nevrologii i psikiatrii im. S.S. Korsakova*. 2017;117(10–12):35–40. (In Russ). DOI: 10.17116/jnevro201711710235-40.

3. Howard J., Trevick S., Younger D.S. Epidemiology of multiple sclerosis. *Neurol Clin.* 2016;34(4):919–939. DOI: 10.1016/j.ncl.2016.06.016. PMID: 27720001.
4. Borodin A.V. Дифференциальная диагностика рассеянного склероза. *Практическая медицина.* 2018;(10):59–63.
5. Siva A. Common clinical and imaging conditions misdiagnosed as multiple sclerosis: a current approach to the differential diagnosis of multiple sclerosis. *Neurol Clin.* 2018;36(1):69–117. DOI: 10.1016/j.ncl.2017.08.014. PMID: 29157405.
6. Solomon A.J., Naismith R.T., Cross A.H. Misdiagnosis of multiple sclerosis: impact of the 2017 McDonald criteria on clinical practice. *Neurology.* 2019;92(1):26–33. DOI: 10.1212/WNL.0000000000006583. PMID: 30381369.
7. Морозов С.П., Владимирский А.В., Кляшторный В.Г. и др. Клинические испытания программного обеспечения на основе интеллектуальных технологий (лучевая диагностика). М., 2019. 33 с.
8. Udupa J.K., Wei L., Samarasekera S., Miki Y. et al. Multiple sclerosis lesion quantification using fuzzy-connectedness principles. *IEEE Trans Med Imaging.* 1997;16(5):598–609. DOI: 10.1109/42.640750. PMID: 9368115.
9. Datta S., Sajja B.R., He R. et al. Segmentation and quantification of black holes in multiple sclerosis. *Neuroimage.* 2006;29(2):467–474. DOI: 10.1016/j.neuroimage.2005.07.042. PMID: 16126416.
10. He R., Sajja B.R., Narayana P.A. Implementation of high-dimensional feature map for segmentation of MR images. *Ann Biomed Eng.* 2005;33(10):1439–1448. DOI: 10.1007/s10439-005-5888-3. PMID: 16240091.
11. Kawa J., Pietka E. Kernelized fuzzy c-means method in fast segmentation of demyelination plaques in multiple sclerosis. *Conf Proc IEEE Eng Med Biol Soc.* 2007;2007:5616–5619. DOI: 10.1109/IEMBS.2007.4353620. PMID: 18003286.
12. Ait-Ali L.S., Prima S., Hellier P. et al. STREM: a robust multidimensional parametric method to segment MS lesions in MRI. *Med Image Comput Assist Interv.* 2005; 8(Pt 1):409–416. DOI: 10.1007/11566465\_51. PMID: 16685872.
13. Raff U., Newman F.D. Automated lesion detection and lesion quantitation in MR images using autoassociative memory. *Med Phys.* 1992;19(1):71–77. DOI: 10.1118/1.596853. PMID: 1620061.
14. Raff U., Newman F.D. Lesion detection in radiologic images using an auto-associative paradigm: preliminary results. *Med Phys.* 1990; 17(5):926–928. DOI: 10.1118/1.596449. PMID: 2233581.
15. Chase H.S., Mitrani L.R., Lu G.G., Fulgieri D.J. Early recognition of multiple sclerosis using natural language processing of the electronic health record. *BMC Med Inform Decis Mak.* 2017;17(1):24. DOI: 10.1186/s12911-017-0418-4.
16. Nelson R.E., Butler J., LaFleur J. et al. Determining multiple sclerosis phenotype from electronic medical records. *J Manag Care Spec Pharm.* 2016;22(12):1377–1382. DOI: 10.18553/jmcp.2016.22.12.1377. PMID: 27882837.
17. Davis M.F., Sriram S., Bush W.S. et al. Automated extraction of clinical traits of multiple sclerosis in electronic medical records. *J Am Med Inform Assoc.* 2013;20(e2):e334–e340. DOI: 10.1136/amiajnl-2013-001999. PMID: 24148554.
18. Wotschel V., Alexander D.C., Kwok P.P. et al. Predicting outcome in clinically isolated syndrome using machine learning. *Neuroimage Clin.* 2014;7:281–287. DOI: 10.1016/j.nicl.2014.11.021. PMID: 25610791.
19. Zhao Y., Healy B.C., Rotstein D. et al. Exploration of machine learning techniques in predicting multiple sclerosis disease course. *PLoS One.* 2017;12(4):e0174866. DOI: 10.1371/journal.pone.0174866. PMID: 28379999.
20. Wu F.Y., Slater J.D., Honig L.S. et al. A neural network design for event-related potential diagnosis. *Comput Biol Med.* 1993;23(3):251–264. DOI: 10.1016/0010-4825(93)90024-u. PMID: 8334865.
21. Ahmadi A., Davoudi S., Daliri M.R. Computer Aided Diagnosis System for multiple sclerosis disease based on phase to amplitude coupling in covert visual attention. *Comput Methods Programs Biomed.* 2019;169:9–18. DOI: 10.1016/j.cmpb.2018.11.006. PMID: 30638593.
22. Kiiski H., Jollans L., Donnchadha S.Ó. et al. Machine learning EEG to predict cognitive functioning and processing speed over a 2-year period in multiple sclerosis patients and controls. *Brain Topogr.* 2018;31(3):346–363. DOI: 10.1007/s10548-018-0620-4. PMID: 29380079.
23. Liberati A., Altman D.G., Tetzlaff J. et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol.* 2009;62(10):e1–e34. DOI: 10.1016/j.jclinepi.2009.06.006. PMID: 19631507.
24. Wang S.H., Tang C., Sun J. et al. Multiple sclerosis identification by 14-layer convolutional neural network with batch normalization, dropout, and stochastic pooling. *Front Neurosci.* 2018;12:818. DOI: 10.3389/fnins.2018.00818. PMID: 30467462.
25. Zurita M., Montalba C., Labbé T. et al. Characterization of relapsing-remitting multiple sclerosis patients using support vector machine classifications of functional and diffusion MRI data. *Neuroimage Clin.* 2018;20:724–730. DOI: 10.1016/j.nicl.2018.09.002. PMID: 30238916.
26. Yoo Y., Tang L.Y.W., Brosch T. et al. Deep learning of joint myelin and T1w MRI features in normal-appearing brain tissue to distinguish between multiple sclerosis patients and healthy controls. *Neuroimage Clin.* 2017;17:169–178. DOI: 10.1016/j.nicl.2017.10.015. PMID: 29071211.
27. Nakamura K., Fisher E. Segmentation of brain magnetic resonance images for measurement of gray matter atrophy in multiple sclerosis patients. *Neuro-*
3. Howard J., Trevick S., Younger D.S. Epidemiology of multiple sclerosis. *Neurol Clin.* 2016;34(4):919–939. DOI: 10.1016/j.ncl.2016.06.016. PMID: 27720001.
4. Borodin A.V. Differential diagnostics of multiple sclerosis. *Prakticheskaya meditsina.* 2018;(10):59–63. (In Russ).
5. Siva A. Common clinical and imaging conditions misdiagnosed as multiple sclerosis: a current approach to the differential diagnosis of multiple sclerosis. *Neurol Clin.* 2018;36(1):69–117. DOI: 10.1016/j.ncl.2017.08.014. PMID: 29157405.
6. Solomon A.J., Naismith R.T., Cross A.H. Misdiagnosis of multiple sclerosis: impact of the 2017 McDonald criteria on clinical practice. *Neurology.* 2019;92(1):26–33. DOI: 10.1212/WNL.0000000000006583. PMID: 30381369.
7. Morozov S.P., Vladzimirsky A.V., Klyashorny V.G. et al. Clinical studies of intelligence technology-based software (radiology). Moscow, 2019. 33 p. (In Russ).
8. Udupa J.K., Wei L., Samarasekera S., Miki Y. et al. Multiple sclerosis lesion quantification using fuzzy-connectedness principles. *IEEE Trans Med Imaging.* 1997;16(5):598–609. DOI: 10.1109/42.640750. PMID: 9368115.
9. Datta S., Sajja B.R., He R. et al. Segmentation and quantification of black holes in multiple sclerosis. *Neuroimage.* 2006;29(2):467–474. DOI: 10.1016/j.neuroimage.2005.07.042. PMID: 16126416.
10. He R., Sajja B.R., Narayana P.A. Implementation of high-dimensional feature map for segmentation of MR images. *Ann Biomed Eng.* 2005;33(10):1439–1448. DOI: 10.1007/s10439-005-5888-3. PMID: 16240091.
11. Kawa J., Pietka E. Kernelized fuzzy c-means method in fast segmentation of demyelination plaques in multiple sclerosis. *Conf Proc IEEE Eng Med Biol Soc.* 2007;2007:5616–5619. DOI: 10.1109/IEMBS.2007.4353620. PMID: 18003286.
12. Ait-Ali L.S., Prima S., Hellier P. et al. STREM: a robust multidimensional parametric method to segment MS lesions in MRI. *Med Image Comput Assist Interv.* 2005; 8(Pt 1):409–416. DOI: 10.1007/11566465\_51. PMID: 16685872.
13. Raff U., Newman F.D. Automated lesion detection and lesion quantitation in MR images using autoassociative memory. *Med Phys.* 1992;19(1):71–77. DOI: 10.1118/1.596853. PMID: 1620061.
14. Raff U., Newman F.D. Lesion detection in radiologic images using an auto-associative paradigm: preliminary results. *Med Phys.* 1990; 17(5):926–928. DOI: 10.1118/1.596449. PMID: 2233581.
15. Chase H.S., Mitrani L.R., Lu G.G., Fulgieri D.J. Early recognition of multiple sclerosis using natural language processing of the electronic health record. *BMC Med Inform Decis Mak.* 2017;17(1):24. DOI: 10.1186/s12911-017-0418-4.
16. Nelson R.E., Butler J., LaFleur J. et al. Determining multiple sclerosis phenotype from electronic medical records. *J Manag Care Spec Pharm.* 2016;22(12):1377–1382. DOI: 10.18553/jmcp.2016.22.12.1377. PMID: 27882837.
17. Davis M.F., Sriram S., Bush W.S. et al. Automated extraction of clinical traits of multiple sclerosis in electronic medical records. *J Am Med Inform Assoc.* 2013;20(e2):e334–e340. DOI: 10.1136/amiajnl-2013-001999. PMID: 24148554.
18. Wotschel V., Alexander D.C., Kwok P.P. et al. Predicting outcome in clinically isolated syndrome using machine learning. *Neuroimage Clin.* 2014;7:281–287. DOI: 10.1016/j.nicl.2014.11.021. PMID: 25610791.
19. Zhao Y., Healy B.C., Rotstein D. et al. Exploration of machine learning techniques in predicting multiple sclerosis disease course. *PLoS One.* 2017;12(4):e0174866. DOI: 10.1371/journal.pone.0174866. PMID: 28379999.
20. Wu F.Y., Slater J.D., Honig L.S. et al. A neural network design for event-related potential diagnosis. *Comput Biol Med.* 1993;23(3):251–264. DOI: 10.1016/0010-4825(93)90024-u. PMID: 8334865.
21. Ahmadi A., Davoudi S., Daliri M.R. Computer Aided Diagnosis System for multiple sclerosis disease based on phase to amplitude coupling in covert visual attention. *Comput Methods Programs Biomed.* 2019;169:9–18. DOI: 10.1016/j.cmpb.2018.11.006. PMID: 30638593.
22. Kiiski H., Jollans L., Donnchadha S.Ó. et al. Machine learning EEG to predict cognitive functioning and processing speed over a 2-year period in multiple sclerosis patients and controls. *Brain Topogr.* 2018;31(3):346–363. DOI: 10.1007/s10548-018-0620-4. PMID: 29380079.
23. Liberati A., Altman D.G., Tetzlaff J. et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol.* 2009;62(10):e1–e34. DOI: 10.1016/j.jclinepi.2009.06.006. PMID: 19631507.
24. Wang S.H., Tang C., Sun J. et al. Multiple sclerosis identification by 14-layer convolutional neural network with batch normalization, dropout, and stochastic pooling. *Front Neurosci.* 2018;12:818. DOI: 10.3389/fnins.2018.00818. PMID: 30467462.
25. Zurita M., Montalba C., Labbé T. et al. Characterization of relapsing-remitting multiple sclerosis patients using support vector machine classifications of functional and diffusion MRI data. *Neuroimage Clin.* 2018;20:724–730. DOI: 10.1016/j.nicl.2018.09.002. PMID: 30238916.
26. Yoo Y., Tang L.Y.W., Brosch T. et al. Deep learning of joint myelin and T1w MRI features in normal-appearing brain tissue to distinguish between multiple sclerosis patients and healthy controls. *Neuroimage Clin.* 2017;17:169–178. DOI: 10.1016/j.nicl.2017.10.015. PMID: 29071211.
27. Nakamura K., Fisher E. Segmentation of brain magnetic resonance images for measurement of gray matter atrophy in multiple sclerosis patients. *Neuro-*

- image*. 2009;44(3):769–776. DOI: 10.1016/j.neuroimage.2008.09.059. PMID: 19007895.
28. Hackmack K., Paul F., Weygandt M. et al. Alzheimer's disease neuroimaging initiative. Multi-scale classification of disease using structural MRI and wavelet transform. *Neuroimage*. 2012;62(1):48–58. DOI: 10.1016/j.neuroimage.2012.05.022. PMID: 22609452.
29. Wei W., Poirion E., Bodini B. et al. Predicting PET-derived demyelination from multimodal MRI using sketcher-refiner adversarial training for multiple sclerosis. *Med Image Anal*. 2019;58:101546. DOI: 10.1016/j.media.2019.101546. PMID: 31499318.
30. Theoharakis P., Glotsos D., Kalatzis I. et al. Pattern recognition system for the discrimination of multiple sclerosis from cerebral microangiopathy lesions based on texture analysis of magnetic resonance images. *Magn Reson Imaging*. 2009;27(3):417–422. DOI: 10.1016/j.mri.2008.07.014. PMID: 18786795.
31. Crimi A., Commowick O., Maarouf A. et al. Predictive value of imaging markers at multiple sclerosis disease onset based on gadolinium- and USPIO-enhanced MRI and machine learning. *PLoS One*. 2014;9(4):e93024. DOI: 10.1371/journal.pone.0093024. PMID: 24691080.
32. Aslani S., Dayan M., Storelli L. et al. Multi-branch convolutional neural network for multiple sclerosis lesion segmentation. *Neuroimage*. 2019;196:1–15. DOI: 10.1016/j.neuroimage.2019.03.068. PMID: 30953833.
33. McKinley R., Wepfer R., Grunder L. et al. Automatic detection of lesion load change in Multiple Sclerosis using convolutional neural networks with segmentation confidence. *Neuroimage Clin*. 2020;25:102104. DOI: 10.1016/j.nicl.2019.102104. PMID: 31927500.
34. Jain S., Ribbens A., Sima D.M. et al. D. Two time point MS lesion segmentation in brain MRI: an expectation-maximization framework. *Front Neurosci*. 2016;10:576. DOI: 10.3389/fnins.2016.00576. PMID: 28066162.
35. Cabezas M., Oliver A., Valverde S. et al. BOOST: a supervised approach for multiple sclerosis lesion segmentation. *J Neurosci Methods*. 2014;237:108–117. DOI: 10.1016/j.jneumeth.2014.08.024. PMID: 25194638.
36. Gabr R.E., Coronado I., Robinson M. et al. Brain and lesion segmentation in multiple sclerosis using fully convolutional neural networks: A large-scale study. *Mult Scler*. 2020;26(10): 1217–1226. DOI: 10.1177/1352458519856843. PMID: 31190607.
37. Brosch T., Tang L.Y., Youngjin Yoo et al. Deep 3D convolutional encoder networks with shortcuts for multiscale feature integration applied to multiple sclerosis lesion segmentation. *IEEE Trans Med Imaging*. 2016;35(5):1229–1239. DOI: 10.1109/TMI.2016.2528821. PMID: 26886978.
38. Valverde S., Cabezas M., Roura E. et al. Improving automated multiple sclerosis lesion segmentation with a cascaded 3D convolutional neural network approach. *Neuroimage*. 2017;155:159–168. DOI: 10.1016/j.neuroimage.2017.04.034. PMID: 28435096.
39. Bossuyt P.M., Reitsma J.B., Bruns D.E. et al. STARD 2015: An Updated List of Essential Items for Reporting Diagnostic Accuracy Studies. *Radiology*. 2015;277(3):826–832. DOI: 10.1148/radiol.2015151516. PMID: 26509226.
40. Yamamoto D., Arimura H., Kakeda S. et al. Computer-aided detection of multiple sclerosis lesions in brain magnetic resonance images: false positive reduction scheme consisted of rule-based, level set method, and support vector machine. *Comput Med Imaging Graph*. 2010;34(5):404–413. DOI: 10.1016/j.compmedimag.2010.02.001. PMID: 20189353.
41. Goldberg-Zimring D., Achiron A., Miron S. et al. Automated detection and characterization of multiple sclerosis lesions in brain MR images. *Magn Reson Imaging*. 1998;16(3):311–318. DOI: 10.1016/s0730-725x(97)00300-7. PMID: 9621972.
42. Bendfeldt K., Klöppel S., Nichols T.E. et al. Multivariate pattern classification of gray matter pathology in multiple sclerosis. *Neuroimage*. 2012;60(1):400–408. DOI: 10.1016/j.neuroimage.2011.12.070. PMID: 22245259.
43. Salem M., Cabezas M., Valverde S. et al. A supervised framework with intensity subtraction and deformation field features for the detection of new T2-w lesions in multiple sclerosis. *Neuroimage Clin*. 2017;17:607–615. DOI: 10.1016/j.nicl.2017.11.015. PMID: 29234597.
44. Boudraa A.O., Dehak S.M., Zhu Y.M. et al. Automated segmentation of multiple sclerosis lesions in multispectral MR imaging using fuzzy clustering. *Comput Biol Med*. 2000;30(1):23–40. DOI: 10.1016/s0010-4825(99)00019-0. PMID: 10695813.
45. Geremia E., Menze B.H., Clatz O. et al. Spatial decision forests for MS lesion segmentation in multi-channel MR images. *Med Image Comput Comput Assist Interv*. 2010;13(Pt 1):111–118. DOI: 10.1007/978-3-642-15705-9\_14. PMID: 20879221.
46. Shiee N., Bazin P.L., Ozturk A. et al. A topology-preserving approach to the segmentation of brain images with multiple sclerosis lesions. *Neuroimage*. 2010;49(2):1524–1535. DOI: 10.1016/j.neuroimage.2009.09.005. PMID: 19766196.
47. Garcia-Lorenzo D., Lecoer J., Arnold D.L. et al. Multiple sclerosis lesion segmentation using an automatic multimodal graph cuts. *Med Image Comput Comput Assist Interv*. 2009;12(Pt 2):584–591. DOI: 10.1007/978-3-642-04271-3\_71. PMID: 20426159.
48. Salem M., Cabezas M., Valverde S. et al. A supervised framework with intensity subtraction and deformation field features for the detection of new T2-w lesions in multiple sclerosis. *Neuroimage Clin*. 2017;17:607–615. DOI: 10.1016/j.nicl.2017.11.015. PMID: 29234597.
- image*. 2009;44(3):769–776. DOI: 10.1016/j.neuroimage.2008.09.059. PMID: 19007895.
28. Hackmack K., Paul F., Weygandt M. et al. Alzheimer's disease neuroimaging initiative. Multi-scale classification of disease using structural MRI and wavelet transform. *Neuroimage*. 2012;62(1):48–58. DOI: 10.1016/j.neuroimage.2012.05.022. PMID: 22609452.
29. Wei W., Poirion E., Bodini B. et al. Predicting PET-derived demyelination from multimodal MRI using sketcher-refiner adversarial training for multiple sclerosis. *Med Image Anal*. 2019;58:101546. DOI: 10.1016/j.media.2019.101546. PMID: 31499318.
30. Theoharakis P., Glotsos D., Kalatzis I. et al. Pattern recognition system for the discrimination of multiple sclerosis from cerebral microangiopathy lesions based on texture analysis of magnetic resonance images. *Magn Reson Imaging*. 2009;27(3):417–422. DOI: 10.1016/j.mri.2008.07.014. PMID: 18786795.
31. Crimi A., Commowick O., Maarouf A. et al. Predictive value of imaging markers at multiple sclerosis disease onset based on gadolinium- and USPIO-enhanced MRI and machine learning. *PLoS One*. 2014;9(4):e93024. DOI: 10.1371/journal.pone.0093024. PMID: 24691080.
32. Aslani S., Dayan M., Storelli L. et al. Multi-branch convolutional neural network for multiple sclerosis lesion segmentation. *Neuroimage*. 2019;196:1–15. DOI: 10.1016/j.neuroimage.2019.03.068. PMID: 30953833.
33. McKinley R., Wepfer R., Grunder L. et al. Automatic detection of lesion load change in Multiple Sclerosis using convolutional neural networks with segmentation confidence. *Neuroimage Clin*. 2020;25:102104. DOI: 10.1016/j.nicl.2019.102104. PMID: 31927500.
34. Jain S., Ribbens A., Sima D.M. et al. D. Two time point MS lesion segmentation in brain MRI: an expectation-maximization framework. *Front Neurosci*. 2016;10:576. DOI: 10.3389/fnins.2016.00576. PMID: 28066162.
35. Cabezas M., Oliver A., Valverde S. et al. BOOST: a supervised approach for multiple sclerosis lesion segmentation. *J Neurosci Methods*. 2014;237:108–117. DOI: 10.1016/j.jneumeth.2014.08.024. PMID: 25194638.
36. Gabr R.E., Coronado I., Robinson M. et al. Brain and lesion segmentation in multiple sclerosis using fully convolutional neural networks: A large-scale study. *Mult Scler*. 2020;26(10): 1217–1226. DOI: 10.1177/1352458519856843. PMID: 31190607.
37. Brosch T., Tang L.Y., Youngjin Yoo et al. Deep 3D convolutional encoder networks with shortcuts for multiscale feature integration applied to multiple sclerosis lesion segmentation. *IEEE Trans Med Imaging*. 2016;35(5):1229–1239. DOI: 10.1109/TMI.2016.2528821. PMID: 26886978.
38. Valverde S., Cabezas M., Roura E. et al. Improving automated multiple sclerosis lesion segmentation with a cascaded 3D convolutional neural network approach. *Neuroimage*. 2017;155:159–168. DOI: 10.1016/j.neuroimage.2017.04.034. PMID: 28435096.
39. Bossuyt P.M., Reitsma J.B., Bruns D.E. et al. STARD 2015: An Updated List of Essential Items for Reporting Diagnostic Accuracy Studies. *Radiology*. 2015;277(3):826–832. DOI: 10.1148/radiol.2015151516. PMID: 26509226.
40. Yamamoto D., Arimura H., Kakeda S. et al. Computer-aided detection of multiple sclerosis lesions in brain magnetic resonance images: false positive reduction scheme consisted of rule-based, level set method, and support vector machine. *Comput Med Imaging Graph*. 2010;34(5):404–413. DOI: 10.1016/j.compmedimag.2010.02.001. PMID: 20189353.
41. Goldberg-Zimring D., Achiron A., Miron S. et al. Automated detection and characterization of multiple sclerosis lesions in brain MR images. *Magn Reson Imaging*. 1998;16(3):311–318. DOI: 10.1016/s0730-725x(97)00300-7. PMID: 9621972.
42. Bendfeldt K., Klöppel S., Nichols T.E. et al. Multivariate pattern classification of gray matter pathology in multiple sclerosis. *Neuroimage*. 2012;60(1):400–408. DOI: 10.1016/j.neuroimage.2011.12.070. PMID: 22245259.
43. Salem M., Cabezas M., Valverde S. et al. A supervised framework with intensity subtraction and deformation field features for the detection of new T2-w lesions in multiple sclerosis. *Neuroimage Clin*. 2017;17:607–615. DOI: 10.1016/j.nicl.2017.11.015. PMID: 29234597.
44. Boudraa A.O., Dehak S.M., Zhu Y.M. et al. Automated segmentation of multiple sclerosis lesions in multispectral MR imaging using fuzzy clustering. *Comput Biol Med*. 2000;30(1):23–40. DOI: 10.1016/s0010-4825(99)00019-0. PMID: 10695813.
45. Geremia E., Menze B.H., Clatz O. et al. Spatial decision forests for MS lesion segmentation in multi-channel MR images. *Med Image Comput Comput Assist Interv*. 2010;13(Pt 1):111–118. DOI: 10.1007/978-3-642-15705-9\_14. PMID: 20879221.
46. Shiee N., Bazin P.L., Ozturk A. et al. A topology-preserving approach to the segmentation of brain images with multiple sclerosis lesions. *Neuroimage*. 2010;49(2):1524–1535. DOI: 10.1016/j.neuroimage.2009.09.005. PMID: 19766196.
47. Garcia-Lorenzo D., Lecoer J., Arnold D.L. et al. Multiple sclerosis lesion segmentation using an automatic multimodal graph cuts. *Med Image Comput Comput Assist Interv*. 2009;12(Pt 2):584–591. DOI: 10.1007/978-3-642-04271-3\_71. PMID: 20426159.
48. Salem M., Cabezas M., Valverde S. et al. A supervised framework with intensity subtraction and deformation field features for the detection of new T2-w lesions in multiple sclerosis. *Neuroimage Clin*. 2017;17:607–615. DOI: 10.1016/j.nicl.2017.11.015. PMID: 29234597.

49. Khayati R., Vafadust M., Towhidkhal F. et al. A novel method for automatic determination of different stages of multiple sclerosis lesions in brain MR FLAIR images. *Comput Med Imaging Graph.* 2008;32(2):124–133. PMID: 18055174.
50. Weiss N., Rueckert D., Rao A. Multiple sclerosis lesion segmentation using dictionary learning and sparse coding. *Med Image Comput Comput Assist Interv.* 2013;16(Pt 1):735–742. DOI: 10.1007/978-3-642-40811-3\_92. PMID: 24505733.
51. Gao J., Li C., Feng C. et al. Non-locally regularized segmentation of multiple sclerosis lesion from multi-channel MRI data. *Magn Reson Imaging.* 2014;32(8):1058–1066. DOI: 10.1016/j.mri.2014.03.006. PMID: 24948583.
52. Bijar A., Khayati R., Peñalver Benavent A. Increasing the contrast of the brain MR FLAIR images using fuzzy membership functions and structural similarity indices in order to segment MS lesions. *PLoS One.* 2013;8(6):e65469. DOI: 10.1371/journal.pone.0065469. PMID: 23799015.
53. Cerasa A., Bilotta E., Augimeri A. et al. A Cellular Neural Network methodology for the automated segmentation of multiple sclerosis lesions. *J Neurosci Methods.* 2012;203(1):193–199. DOI: 10.1016/j.jneumeth.2011.08.047. PMID: 21920384.
54. Aymerich F.X., Sobrevilla P., Montseny E. et al. Fuzzy approach toward reducing false positives in the detection of small multiple sclerosis lesions in magnetic resonance images. *Conf Proc IEEE Eng Med Biol Soc.* 2011;2011:5694–5697. DOI: 10.1109/IEMBS.2011.6091378. PMID: 22255632.
55. Kuwazuru J., Arimura H., Kakeda S. et al. Automated detection of multiple sclerosis candidate regions in MR images: false-positive removal with use of an ANN-controlled level-set method. *Radiol Phys Technol.* 2012;5(1):105–113. DOI: 10.1007/s12194-011-0141-2. PMID: 22139608.
56. Kanber B., Nachev P., Barkhof F. et al. High-dimensional detection of imaging response to treatment in multiple sclerosis. *NPJ Digit Med.* 2019;2:49. DOI: 10.1038/s41746-019-0127-8. PMID: 31304395.
57. Freire P.G.L., Ferrari R.J. Multiple sclerosis lesion enhancement and white matter region estimation using hyperintensities in FLAIR images. *arXiv.* 2018:1807.09619.
58. Feng Y., Pan H., Meyer C. et al. A self-adaptive network for multiple sclerosis lesion segmentation from multi-contrast MRI with various imaging protocols. *arXiv.* 2018:1811.07491.
59. Kazancli E., Prchkovska V., Rodrigues P. et al. Multiple sclerosis lesion segmentation using improved convolutional neural networks. In: *Proceedings of the 13<sup>th</sup> International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2018)*. 2018; 4; VISAPP: 260–269. DOI: 10.5220/0006540902600269.
60. Roy S., Butman J.A., Reich D.S. et al. Multiple sclerosis lesion segmentation from brain MRI via fully convolutional neural networks. *arXiv.* 2018:1803.09172v1.
61. Saccà V., Sarica A., Novellino F. et al. Evaluation of machine learning algorithms performance for the prediction of early multiple sclerosis from resting-state FMRI connectivity data. *Brain Imaging Behav.* 2019;13(4):1103–1114. DOI: 10.1007/s11682-018-9926-9. PMID: 29992392.
62. Zhang H., Zhang J., Zhang Q. et al. RSANet: Recurrent Slice-wise Attention Network for multiple sclerosis lesion segmentation. *arXiv;* 2020:2002.12470v1.
63. Gheshlaghi S.H., Ranjbar A., Suratgar A.A. et al. A superpixel segmentation based technique for multiple sclerosis lesion detection. *arXiv;* 2020:1907.03109v1.
49. Khayati R., Vafadust M., Towhidkhal F. et al. A novel method for automatic determination of different stages of multiple sclerosis lesions in brain MR FLAIR images. *Comput Med Imaging Graph.* 2008;32(2):124–133. PMID: 18055174.
50. Weiss N., Rueckert D., Rao A. Multiple sclerosis lesion segmentation using dictionary learning and sparse coding. *Med Image Comput Comput Assist Interv.* 2013;16(Pt 1):735–742. DOI: 10.1007/978-3-642-40811-3\_92. PMID: 24505733.
51. Gao J., Li C., Feng C. et al. Non-locally regularized segmentation of multiple sclerosis lesion from multi-channel MRI data. *Magn Reson Imaging.* 2014;32(8):1058–1066. DOI: 10.1016/j.mri.2014.03.006. PMID: 24948583.
52. Bijar A., Khayati R., Peñalver Benavent A. Increasing the contrast of the brain MR FLAIR images using fuzzy membership functions and structural similarity indices in order to segment MS lesions. *PLoS One.* 2013;8(6):e65469. DOI: 10.1371/journal.pone.0065469. PMID: 23799015.
53. Cerasa A., Bilotta E., Augimeri A. et al. A Cellular Neural Network methodology for the automated segmentation of multiple sclerosis lesions. *J Neurosci Methods.* 2012;203(1):193–199. DOI: 10.1016/j.jneumeth.2011.08.047. PMID: 21920384.
54. Aymerich F.X., Sobrevilla P., Montseny E. et al. Fuzzy approach toward reducing false positives in the detection of small multiple sclerosis lesions in magnetic resonance images. *Conf Proc IEEE Eng Med Biol Soc.* 2011;2011:5694–5697. DOI: 10.1109/IEMBS.2011.6091378. PMID: 22255632.
55. Kuwazuru J., Arimura H., Kakeda S. et al. Automated detection of multiple sclerosis candidate regions in MR images: false-positive removal with use of an ANN-controlled level-set method. *Radiol Phys Technol.* 2012;5(1):105–113. DOI: 10.1007/s12194-011-0141-2. PMID: 22139608.
56. Kanber B., Nachev P., Barkhof F. et al. High-dimensional detection of imaging response to treatment in multiple sclerosis. *NPJ Digit Med.* 2019;2:49. DOI: 10.1038/s41746-019-0127-8. PMID: 31304395.
57. Freire P.G.L., Ferrari R.J. Multiple sclerosis lesion enhancement and white matter region estimation using hyperintensities in FLAIR images. *arXiv.* 2018:1807.09619.
58. Feng Y., Pan H., Meyer C. et al. A self-adaptive network for multiple sclerosis lesion segmentation from multi-contrast MRI with various imaging protocols. *arXiv.* 2018:1811.07491.
59. Kazancli E., Prchkovska V., Rodrigues P. et al. Multiple sclerosis lesion segmentation using improved convolutional neural networks. In: *Proceedings of the 13<sup>th</sup> International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2018)*. 2018; 4; VISAPP: 260–269. DOI: 10.5220/0006540902600269.
60. Roy S., Butman J.A., Reich D.S. et al. Multiple sclerosis lesion segmentation from brain MRI via fully convolutional neural networks. *arXiv.* 2018:1803.09172v1.
61. Saccà V., Sarica A., Novellino F. et al. Evaluation of machine learning algorithms performance for the prediction of early multiple sclerosis from resting-state FMRI connectivity data. *Brain Imaging Behav.* 2019;13(4):1103–1114. DOI: 10.1007/s11682-018-9926-9. PMID: 29992392.
62. Zhang H., Zhang J., Zhang Q. et al. RSANet: Recurrent Slice-wise Attention Network for multiple sclerosis lesion segmentation. *arXiv;* 2020:2002.12470v1.
63. Gheshlaghi S.H., Ranjbar A., Suratgar A.A. et al. A superpixel segmentation based technique for multiple sclerosis lesion detection. *arXiv;* 2020:1907.03109v1.

## Информация об авторах

*Черняева Галина Николаевна* — м.н.с. ГБУЗ г. Москвы «Научно-практический клинический центр диагностики и телемедицинских технологий», Москва, Россия

*Морозов Сергей Павлович* — д.м.н., проф., директор ГБУЗ г. Москвы «Научно-практический клинический центр диагностики и телемедицинских технологий», Москва, Россия, [orcid.org/0000-0001-6545-6170](https://orcid.org/0000-0001-6545-6170)

*Владимирский Антон Вячеславович* — д.м.н., заместитель директора по научной работе ГБУЗ г. Москвы «Научно-практический клинический центр диагностики и телемедицинских технологий», Москва, Россия; профессор каф. информационных и интернет технологий ФГАОУ ВО «ПМГМУ им. И.М. Сеченова», Москва, Россия, [orcid.org/0000-0002-2990-7736](https://orcid.org/0000-0002-2990-7736)

**Вклад авторов.** *Черняева Г.Н.* — создание концепции исследования, курирование данных, анализ данных, подготовка черновика рукописи; *Морозов С.П.* — руководство научно-исследовательской работой, доработка и редактирование рукописи; *Владимирский А.В.* — курирование данных, анализ данных, подготовка черновика рукописи.

## Information about the authors

*Galina N. Chernyaeva* — junior researcher, Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies, Moscow, Russia

*Sergey P. Morozov* — D. Sci. (Med), Prof., Director, Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies, Moscow, Russia, [orcid.org/0000-0001-6545-6170](https://orcid.org/0000-0001-6545-6170)

*Anton V. Vladimirovsky* — D. Sci. (Med), Deputy Director for R&D, Moscow Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies, Moscow, Russia; Professor, Chair of information and internet technologies, I.M. Sechenov First Moscow State Medical University, Moscow, Russia, [orcid.org/0000-0002-2990-7736](https://orcid.org/0000-0002-2990-7736)

**Author contribution.** *Chernyaeva G.N.* — creation of a research concept, data curation, data analysis, preparation of a draft manuscript; *Morozov S.P.* — management of research work, revision and editing of the manuscript; *Vladimirovsky A.V.* — data curation, data analysis, preparation of a draft manuscript.