

Саидов Комрон Жуманазарович¹ <https://orcid.org/0000-0003-0605-5080>

1. Специализированный научно-практический центр
нейрохирургии и нейрореабилитации при Самаркандском
государственном медицинском университете
Самарканд, Узбекистан

АНАЛИЗ РЕЗУЛЬТАТОВ ХИРУРГИЧЕСКОГО ЛЕЧЕНИЯ ГРЫЖИ ПОЯСНИЧНОГО ДИСКА С ИСПОЛЬЗОВАНИЕМ КЛИНИЧЕСКИХ И РАДИОМИЧЕСКИХ ХАРАКТЕРИСТИК

Аннотация

Введение. Боль в пояснице — широко распространенный симптом и основная причина инвалидности во всем мире. Хотя различные дегенеративные изменения, выявляемые при магнитно-резонансной томографии (МРТ), связывают с болью в пояснице и грыжей межпозвоночного диска, ни одно из них нельзя считать патогномоничным для этого состояния, учитывая высокую распространенность патологических изменений у бессимптомных лиц. Тем не менее, существует недостаток знаний о том, могут ли радиомические признаки на изображениях МРТ в сочетании с клиническими признаками быть полезны для моделирования прогнозирования успеха лечения. Целью данного исследования было изучение потенциала анализа радиомических признаков в сочетании с клиническими признаками и методами искусственного интеллекта (машинное обучение/глубокое обучение) для выявления предикторов МРТ для прогнозирования результатов после операции по поводу грыжи межпозвоночного диска поясничного отдела позвоночника.

Материалы и методы. В исследование были включены 172 пациента, перенесших дискэктомию по поводу грыжи межпозвоночного диска, у которых были выполнены предоперационные МРТ-исследования в режиме T2. Извлеченные клинические характеристики включали пол, возраст, потребление алкоголя и никотина, тип страхования, продолжительность пребывания в стационаре, осложнения, время операции, оценку по шкале ASA, предоперационный уровень С-реактивного белка (СРБ), хирургическую технику (микрохирургическую или эндоскопическую) и информацию об опыте хирурга (количество лет опыта работы с данной хирургической техникой и количество операций, выполненных на момент операции). В данном исследовании использовался полуавтоматический алгоритм объемной сегментации с помощью роста областей для сегментации грыж межпозвоночных дисков. Кроме того, из изображений компьютерной магнитно-резонансной томографии (МРТ) были извлечены 3D-радиомические признаки, характеризующие фенотипические различия на основе интенсивности, формы и текстуры. Выбранные признаки, выявленные с помощью анализа важности признаков, были использованы как для моделей машинного обучения, так и для моделей глубокого обучения (17 моделей).

Результаты. Средняя точность всех моделей при обучении и тестировании в объединенном наборе признаков составила $93,31 \pm 4,96$ и $88,17 \pm 2,58$ соответственно. Средняя точность при обучении и тестировании в клиническом наборе признаков составила $91,28 \pm 4,56$ и $87,69 \pm 3,62$ соответственно.

Выводы. Наши результаты указывают на минимальное, но обнаруживаемое улучшение в задачах прогнозирования при включении радиомических признаков. Однако степень этого преимущества следует оценивать с осторожностью, подчеркивая потенциал использования мультимодальных входных данных в будущих моделях прогнозирования.

Ключевые слова: грыжа межпозвонкового диска; поясничный отдел; дискэктомия; радиомика; MPT; T2-взвешенные изображения; машинное обучение; глубокое обучение; искусственный интеллект; прогнозирование исходов; осложнения; продолжительность пребывания (LOS); время операции; AUC; ROC-анализ.

Saidov Komron Jumanazarovich¹

1. Specialized Scientific and Practical Center of Neurosurgery and
Neurorehabilitation, Samarkand State Medical University
Samarkand, Uzbekistan

ANALYSIS OF SURGICAL OUTCOMES OF LUMBAR DISC HERNIATION USING CLINICAL AND RADIOMIC FEATURES

Abstract

Introduction. Low back pain is a highly prevalent symptom and a leading cause of disability worldwide. Although various degenerative changes detected by magnetic resonance imaging (MRI) are associated with low back pain and intervertebral disc herniation, none of these findings can be considered pathognomonic, given the high prevalence of similar abnormalities in asymptomatic individuals. At the same time, there is limited evidence on whether radiomic features extracted from MRI, when combined with clinical variables, can be useful for predictive modeling of treatment outcomes. The aim of this study was to investigate the potential of radiomic feature analysis combined with clinical characteristics and artificial intelligence methods (machine learning/deep learning) to identify MRI-based predictors of postoperative outcomes after lumbar disc herniation surgery.

Materials and Methods. The study included 172 patients who underwent discectomy for lumbar intervertebral disc herniation and had preoperative T2-weighted MRI scans. Extracted clinical characteristics included sex, age, alcohol and nicotine consumption, insurance status, length of hospital stay, postoperative complications, operative time, ASA score, preoperative C-reactive protein (CRP) level, surgical technique (microsurgical or endoscopic), and surgeon experience (years of experience with the respective technique and number of procedures performed at the time of surgery). A semi-automatic volumetric region-growing segmentation algorithm was used to delineate disc herniations. In addition, three-dimensional radiomic features describing intensity, shape, and texture were extracted from MRI images. Features selected through feature importance analysis were subsequently used to train machine learning and deep learning models (17 models).

Results. The mean training and testing accuracy across all models using the combined feature set was $93.31 \pm 4.96\%$ and $88.17 \pm 2.58\%$, respectively. For models based on clinical features alone, the mean training and testing accuracy was $91.28 \pm 4.56\%$ and $87.69 \pm 3.62\%$, respectively.

Conclusions. Our results demonstrate a minimal but detectable improvement in predictive performance when radiomic features are included. However, the magnitude of this advantage should be interpreted with caution, while highlighting the potential of multimodal input data for future predictive modeling approaches.

Key words: lumbar disc herniation; discectomy; radiomics; MRI; T2-weighted imaging; machine learning; deep learning; artificial intelligence; outcome prediction; postoperative complications; length of stay (LOS); operative time; AUC; ROC analysis.

Saidov Komron Jumanarovich¹

1. Samarqand davlat tibbiyot universiteti huzuridagi
Ixtisoslashtirilgan ilmiy-amaliy neyroxirurgiya va neyroreabilitatsiya
markazi
Samarqand, O'zbekiston

BEL UMURTQA ORALIG'I DISKI CHURRASI JARROHLIK DAVOLASH NATIJALARINI KLINIK VA RADIOMIK XUSUSIYATLAR ASOSIDA TAHLIL QILISH Annotatsiya.

Kirish. Bel sohasidagi og'riq butun dunyo bo'yicha keng tarqalgan simptom bo'lib, nogironlikning asosiy sabablaridan biri hisoblanadi. Magnit-rezonans tomografiya (MRT) yordamida aniqlanadigan turli degenerativ o'zgarishlar bel og'rig'i va umurtqalararo disk churrasi bilan bog'liq bo'lsa-da, asemptomatik shaxslarda ham bunday o'zgarishlarning yuqori uchrash darajasi tufayli ularning hech biri patognomonik belgi sifatida qaralishi mumkin emas. Shu bilan birga, MRT tasvirlaridan olingan radiomik xususiyatlarning klinik ko'rsatkichlar bilan birgalikda davolash natijalarini bashoratlashda foydaliligi yetarli darajada o'rganilmagan. Ushbu tadqiqotning maqsadi bel umurtqa oraliq diski churrasini jarrohlik yo'li bilan davolashdan keyingi natijalarni bashoratlash uchun MRT asosidagi radiomik xususiyatlar, klinik ko'rsatkichlar va sun'iy intellekt (mashinali o'rganish/chuqur o'rganish) usullarining salohiyatini baholashdan iborat.

Materiallar va usullar. Tadqiqotga bel umurtqa oraliq diski churrasi sababli diskektomiya o'tkazilgan va operatsiyadan oldin T2-rejimdagi MRT tekshiruvidan o'tgan 172 nafar bemor kiritildi. Klinik ko'rsatkichlar sifatida jins, yosh, alkohol va nikotin iste'moli, sug'urta turi, statsionarda qolish davomiyligi, operatsiyadan keyingi asoratlari, operatsiya vaqti, ASA bahosi, operatsiyagacha C-reaktiv oqsil (CRP) darajasi, jarrohlik texnikasi (mikroxirurgik yoki endoskopik) hamda jarroh tajribasi (mazkur texnika bo'yicha ish staji va operatsiya vaqtigacha bajarilgan operatsiyalar soni) hisobga olindi. Umurtqalararo disk churrasini segmentatsiya qilish uchun yarim avtomatlashtirilgan hajmiy region-growing algoritmi qo'llanildi. MRT tasvirlaridan intensivlik, shakl va teksturani tavsiflovchi uch o'lchamli radiomik xususiyatlar ajratib olindi. Muhimligi aniqlangan tanlangan xususiyatlar mashinali va chuqur o'rganish modellari (17 ta model)ni qurishda qo'llanildi.

Natijalar. Birlashtirilgan klinik va radiomik xususiyatlar to'plamidan foydalangan holda barcha modellarda o'rtacha o'qitish aniqligi $93,31 \pm 4,96\%$ va test aniqligi $88,17 \pm 2,58\%$ ni tashkil etdi. Faqat klinik xususiyatlarga asoslangan modellarda esa o'rtacha o'qitish aniqligi $91,28 \pm 4,56\%$ va test aniqligi $87,69 \pm 3,62\%$ bo'ldi.

Xulosa. Olingan natijalar radiomik xususiyatlarni qo'shish bashoratlash samaradorligini biroz, ammo sezilarli darajada oshirishini ko'rsatadi. Shu bilan birga, bu ustunlikning miqyosi ehtiyotkorlik bilan talqin qilinishi lozim bo'lib, kelajakdagi bashoratlovchi modellar uchun multimodal ma'lumotlardan foydalanish salohiyatini ta'kidlaydi.

Kalit so'zlar: bel umurtqa oraliq diski churrasi; diskektomiya; radiomika; MRT; T2-vaznli tasvirlar; mashinali o'rganish; chuqur o'rganish; sun'iy intellekt; natijani prognozlash; asoratlar; statsionarda qolish (LOS); operatsiya vaqti; AUC; ROC tahlili.

Введение. Грыжа межпозвонкового диска (ГМПД) — распространенное заболевание, поражающее нижнюю часть спины [1]. Первичным местом возникновения ГМПД является межпозвонковый диск (МПД), который характеризуется смещением пульпозного ядра через разрыв фиброзного кольца. Грыжа МПД может привести к сдавливанию корешка спинномозгового нерва, вызывая радикулопатию, которая часто сопровождается сильной болью, онемением и мышечной слабостью в нижних конечностях [2]. Несмотря на значительные успехи в диагностике и лечении ГМПД, ее этиология и патогенез остаются недостаточно изученными [3]. В развитии ГМПД участвуют различные факторы, включая генетику, возраст, профессиональные и связанные с образом жизни факторы, а также биомеханику позвоночника [4]. Диагноз ГМПД в основном основывается на клиническом обследовании, включающем анамнез и физикальное обследование, при этом магнитно-резонансная томография (МРТ) является предпочтительным методом визуализации для подтверждения диагноза [5]. Хотя на начальном этапе рекомендуется консервативное лечение, хирургическое вмешательство может быть рассмотрено в случаях постоянной боли или неврологических нарушений. Несмотря на высокую эффективность хирургического вмешательства, послеоперационные результаты могут значительно различаться в зависимости от ряда факторов, таких как факторы, связанные с пациентом, хирургическая техника или опыт хирургов, а также тяжесть состояния [6]. Радиомика, быстро развивающаяся область, включает в себя извлечение количественных характеристик из медицинских изображений с использованием передовых методов визуализации, таких как магнитно-резонансная томография (МРТ) [7]. Радиомические характеристики показали свою перспективность в прогнозировании результатов лечения и прогноза при нескольких типах рака [8]. Однако существует недостаток исследований, изучающих потенциал сочетания радиомических характеристик с клиническими переменными для прогнозирования послеоперационных результатов у пациентов с грыжей межпозвонкового диска. Эти радиомические характеристики могут быть дополнительно проанализированы в сочетании с клиническими переменными с использованием методов искусственного интеллекта, а именно машинного обучения и глубокого обучения [9]. Машинное обучение и глубокое обучение являются подобластями искусственного интеллекта (ИИ), которые включают разработку алгоритмов для обучения и прогнозирования на основе данных. Хотя оба подхода имеют сходства, между ними существуют заметные различия [10]: модели машинного обучения обычно используют различные алгоритмические методы, такие как деревья решений, машины опорных векторов и логистическая регрессия, для выявления закономерностей в данных и прогнозирования. Эти модели часто требуют ручной обработки признаков, когда эксперты в предметной области выбирают релевантные признаки из входных данных для эффективного обучения алгоритмов. Глубокое обучение, с другой стороны, представляет собой специализированное подмножество машинного обучения, которое использует искусственные нейронные сети (ИНС) для автоматического обучения и извлечения признаков из необработанных данных без

необходимости ручной обработки признаков [11]. Модели глубокого обучения, такие как сверточные нейронные сети (CNN) и рекуррентные нейронные сети (RNN), могут обрабатывать сложные структуры данных, включая изображения, речь и текст, и продемонстрировали замечательную производительность в таких задачах, как распознавание изображений, обработка естественного языка и распознавание речи [9]. Вкратце, хотя и машинное обучение, и глубокое обучение направлены на создание прогностических моделей на основе данных, модели глубокого обучения специально используют искусственные нейронные сети для автоматического обучения и извлечения признаков, часто достигая превосходных результатов в задачах с многомерными и сложными данными [9]. В этом исследовании мы стремились выяснить, может ли сочетание радиомических признаков, извлеченных из предоперационных МРТ-изображений, с клиническими признаками улучшить прогнозирование результатов после операции по поводу грыжи межпозвоночного диска. Мы предположили, что сочетание радиомических и клинических признаков обеспечит более точное прогнозирование послеоперационных результатов, чем использование только радиомических или только клинических признаков. Для достижения этой цели мы оценили совокупное влияние на время операции, продолжительность пребывания в больнице и частоту осложнений, создав составную переменную интересующего нас результата, что позволило провести всестороннюю и обобщенную оценку результатов лечения пациентов.

Материалы и Методы. Дизайн исследования. Отделение хирургии позвоночника больницы Лоретто-Фрайбург, входящей в состав Университетского медицинского центра Фрайбурга, провело ретроспективное когортное исследование для изучения результатов микрохирургической и эндоскопической интерламинарной хирургической техники у пациентов с грыжей поясничного диска. В исследование были включены последовательно поступающие пациенты, которым были выполнены эти хирургические вмешательства в период с 2016 по 2021 год. Перед проведением этого ретроспективного наблюдательного исследования было получено одобрение местного этического комитета Фрайбурга, Германия [номер: 116/200]. Перед участием в исследовании у каждого пациента было получено информированное письменное согласие. Критерии включения в исследование включали пациентов с грыжей поясничного диска, которым в указанный период были выполнены либо микрохирургические, либо эндоскопические вмешательства. В исследование были включены только пациенты, которым было проведено предоперационное МРТ-исследование. В рамках эндоскопической процедуры использовалась система iLESSYS® (Joimax GmbH, Карлсруэ, Германия). Для фильтрации исходного набора данных мы применили наши критерии исключения после сбора всех данных от пациентов, удовлетворяющих нашим критериям включения. Мы исключили пациентов моложе 18 лет, пациентов с опухолями позвоночника или после спондилодеза, а также тех, кто отказался от использования своих данных для исследований.

Обработка данных. В исследовании использовались данные из системы информации о пациентах, а для псевдонимизации данных применялась команда «encode» в статистическом программном обеспечении Stata версии 15. На основе предыдущих исследований и обзора литературы были определены клинические переменные как потенциально значимые факторы, определяющие клинические исходы (рассматриваемые целевые переменные:

продолжительность пребывания в больнице, время операции и осложнения). Эти клинические характеристики включали хирургическую технику, классификацию физического состояния ASA, демографические данные. Перед процедурой у каждого пациента были получены T2-взвешенные изображения (T2WI). Все участники прошли магнитно-резонансную томографию (MPT) на 3-тесловом сканере (Siemens MAGNETOM Skyra). T2-взвешенные изображения были получены с использованием последовательности Turbo SpinEcho с временем повторения 4500 мс, временем эхо 100 мс и полем зрения 220×220 мм. Размер матрицы был установлен на 384×384, что обеспечило разрешение в плоскости 0,57×0,57 мм. Толщина среза составляла 3 мм. Ширина полосы пропускания для последовательности составляла 180 Гц/пиксель. На основе T2WI были определены области интереса (VOI) в области грыжи межпозвоночного диска, оцененные в сагиттальной плоскости каждого пациента. Использовалось радиомическое расширение программного обеспечения 3D Slicer под названием SlicerRadiomics (V2.10, <http://github.com/Radiomics/SlicerRadiomics>), которое включает информацию о предоперационных уровнях С-реактивного белка. В исследовании также были собраны имена хирургов, выполнявших операции в течение периода исследования, с акцентом на оценку их опыта и количества операций, выполненных с использованием соответствующей хирургической техники на момент операции. Хирурги были сгруппированы по количеству выполненных ими операций: пять хирургов выполнили большинство операций, а остальные были объединены в группу, если они выполнили менее десяти операций. Результаты лечения пациентов отслеживались в течение периода до одного года после операции. Для нашей когорты среднее время наблюдения составило 7±4 месяца. Для извлечения радиомических признаков из VOI использовалась библиотека PyRadiomics (22). Процесс сегментации был выполнен двумя врачами-исследователями с 2 и 3 годами опыта сегментации изображений для алгоритмов ИИ, которые обратились за помощью к опытному радиологу с более чем пятилетним опытом сегментации изображений для приложений искусственного интеллекта. Радиолог осуществлял надзор во время процедуры сегментации и внес вклад в разработку алгоритма сегментации, который был реализован с использованием программной платформы 3D Slicer. Для сегментации грыжи межпозвоночного диска использовался полуавтоматический метод, заключающийся в ручном определении частей сегмента грыжи и получении локальных гистограмм интенсивности. На следующем этапе эти пороговые значения использовались для расширения объемной сегментации диска и прилегающих к нему срезов. Затем с помощью 3D-Slicer из сегментированных данных были извлечены радиомические признаки. Извлеченные радиомические признаки перечислены в дополнительной таблице S1. Признаки включали статистику первого порядка, признаки на основе формы (3D), признаки на основе формы (2D), матрицу совместной встречаемости уровней серого (GLCM), матрицу длин серий уровней серого (GLRLM), матрицу зон размера уровней серого (GLSZM), матрицу разности соседних уровней серого (NGTDM) и матрицу зависимости уровней серого (GLDM). Признаки на основе формы (3D и 2D) включают признаки трехмерного и двухмерного размера и формы VOI. Эти признаки не зависят от распределения интенсивности уровней серого в VOI и, следовательно, вычисляются только на непроизводном изображении и маске. GLCM размером NgXNg

характеризует функцию совместной вероятности второго порядка для заданной области изображения, ограниченной маской.

Благодаря этому процессу создается трехмерное представление объема, что позволяет проводить более всесторонний анализ и визуализацию данных между двумя пикселями на изображении, которые разделены расстоянием δ пикселей вдоль угла θ . GLSZM количественно определяет зоны уровней серого на изображении, где зона относится к группе соединенных вокселей, имеющих одинаковую интенсивность уровня серого. GLRLM количественно определяет последовательности уровней серого, которые представляют собой последовательные пиксели с одинаковым значением уровня серого, представленные их длиной в количестве пикселей. NGTDM количественно определяет отклонение между значением серого и соседним средним значением серого на расстоянии δ . В матрице хранится сумма абсолютных разностей для каждого уровня серого i . Наконец, GLDM количественно определяет зависимости уровней серого на изображении, где зависимость относится к количеству соединенных вокселей на расстоянии δ , которые зависят от центрального вокселя [12, 13]. Для оценки внутри- и межнаблюдательского согласия было случайным образом отобрано тридцать образцов из числа пациентов, участвовавших в исследовании. Для оценки внутринаблюдательной воспроизводимости два эксперта независимо друг от друга дважды в течение одной недели определяли области интереса (VOI). Для определения перекрытия областей интереса мы использовали коэффициент Дайса. Расчет коэффициента Дайса был выполнен с помощью процедуры SimpleITK, работающей на Python. Кроме того, для оценки внутринаблюдательного и межнаблюдательного согласия всех радиомических признаков, полученных из сегментации областей интереса, выполненной двумя экспертами, использовался коэффициент внутриклассовой корреляции (ICC). Согласно предыдущему исследованию [31], радиомические признаки с внутри- и межнаблюдательными ICC равными 0,75 считались надежными радиомическими признаками и были включены в окончательный анализ.

Результаты. В анализ были включены в общей сложности $n=172$ пациента (81 [47,1%] женщин; 91 [52,9%] мужчин). Средний возраст составил $59,19 \pm 16,49$ лет (диапазон: 27–92) (Таблица 1). Большинство пациентов имели предоперационный балл ASA 2 (102; 59,3%) и несчастную страховку (129; 75%). $N=112$ (65,1%) пациентов перенесли микрохирургическую операцию на поясничном межпозвоночном диске, а $n=60$ (34,9%) — эндоскопическую операцию на межпозвоночном диске. Среднее время операции составило $60,56 \pm 38,74$ мин, а средняя продолжительность пребывания в стационаре — $13,09 \pm 8,56$ дней. На момент проведения операции хирурги имели опыт работы $6,43 \pm 4,22$ года (диапазон: 0–16 лет) с использованием соответствующей хирургической техники и выполнили в среднем $330,32 \pm 484,74$ операции (диапазон: 5–1820), что указывает на широкий диапазон различных уровней подготовки хирургов. Поэтому мы учитывали опыт хирургов в дальнейших анализах, чтобы учесть высокую дисперсию в моделировании прогнозирования.

Сравнение результатов популяционных исследований. В таблице 2 представлено сравнение двух целевых групп исследования, сформированных на основе данных о продолжительности пребывания в стационаре, времени операции и осложнениях ($n=152$ [88,4%] нормальные; $n=20$

[11,6%] худшие). Как и ожидалось в соответствии с процедурой формирования групп, наблюдалась значительная разница в

Таблица 1. Описательная статистика когорты. ИМТ: индекс массы тела; ASA: классификация риска Американского общества анестезиологов; СРБ: С-реактивный белок; LOS: продолжительность пребывания в стационаре.

Переменная	Среднее значение $\pm$ стандартное отклонение / Количество (N, %)
Возраст (лет)	59 $\pm$ 16
Секс	
Мужской	91 (52,9%)
Женский	81 (47,1%)
ИМТ (кг/м ²)	28 $\pm$ 6
употребление никотина	
Нет	123 (71,5%)
Да	49 (28,5%)
употребление алкоголя	
Нет	110 (64,0%)
Да	62 (36,0%)
страховой статус	
Частный	43 (25,0%)
Нечастный	129 (75,0%)
оценка ASA	
ACA I	55 (32,0%)
ASA II	102 (59,3%)
ASA III	15 (8,7%)
Предоперационный С-реактивный белок	10 $\pm$ 32
Продолжительность пребывания (дни)	13 $\pm$ 9
Время выполнения операции (мин)	61 $\pm$ 39
Многолетний опыт проведения хирургических операций различного типа.	6 $\pm$ 4
Количество предыдущих операций аналогичного типа	330 $\pm$ 485

Таблица 2. Сравнение переменных исследования между двумя группами целевых результатов (ЦРР). ИМТ: индекс массы тела; ASA: классификация риска Американского общества анестезиологов; СРБ: С-реактивный белок; LOS: продолжительность пребывания в стационаре.

Переменная	Норма (среднее значение $\pm$ стандартное отклонение / N%)	Хуже (среднее значение $\pm$ стандартное отклонение / N%)	p-значение
Возраст (лет)	57 $\pm$ 16	68 $\pm$ 12	0,006

Пол – Мужской	81 (53.3)	10 (50.0)	0,048
Пол – женский	71 (46.7)	10 (50.0)	
ИМТ (кг/м ²)	27 ± 6	29 ± 5	0,561
Никотин – Нет	107 (70,4)	16 (80.0)	0.701
Никотин – Да	45 (29.6)	4 (20.0)	
Алкоголь – Запрещено	100 (65,8)	10 (50.0)	0,812
Алкоголь – Да	52 (34.2)	10 (50.0)	
Страхование – частное	36 (23.7)	7 (35.0)	0,104
Страхование – нечастное	116 (76.3)	13 (65.0)	
АСА I	49 (32.2)	6 (30.0)	0,487
АСА II	91 (59,9)	11 (55.0)	
АСА III	12 (7.9)	3 (15.0)	
Предоперационный С-реактивный белок	8 ± 25	30 ± 65	0,018
Хирургическая техника – микрохирургическая	96 (63.2)	14 (70.0)	0,941
Хирургическая техника – Полноэндоскопическая	56 (36,8)	6 (30.0)	
Продолжительность пребывания (дни)	13 ± 8	21 ± 10	<0.001
Время выполнения операции (мин)	58 ± 38	92 ± 46	<0.001
Осложнений – нет.	120 (78,9)	1 (5.0)	<0.001
Осложнения – Да	32 (21.1)	19 (95.0)	
Многолетний опыт	7 ± 4	6 ± 5	0,214
Количество ранее перенесенных операций	325 ± 480	410 ± 570	0.901

Частота осложнений, продолжительность пребывания в стационаре и время операции различались между двумя группами с худшим исходом ($p < 0,001$), при этом в группе с худшим исходом наблюдались более длительная продолжительность пребывания в стационаре, большее время операции и более высокая частота осложнений. Кроме того, пациенты в группе с худшим исходом были значительно старше ($p = 0,004$) и имели более высокий предоперационный уровень С-реактивного белка, что указывает на предоперационное воспалительное состояние ($p = 0,015$). Мы не наблюдали существенных различий в отношении хирургической техники (микрохирургическая против полностью эндоскопической) и кривой обучения хирурга, выполнявшего операцию.

Прогностическое моделирование с использованием радиомики и методов, основанных на искусственном интеллекте.

На следующем этапе мы оценили, приведет ли сочетание характеристик МРТ с клиническими переменными к лучшей прогностической эффективности, чем использование только клинических переменных. Первоначальный анализ важности признаков показал, что возраст и предоперационный уровень С-реактивного белка (СРБ) являются наиболее важными клиническими признаками, тогда как наиболее важными радиомическими признаками являются наборы признаков GLCM, статистики первого порядка и NGTDM. Результаты работы алгоритмов с наивысшей эффективностью для комбинированных радиомических и клинических переменных, а также для алгоритмов, использующих только клинические переменные, показаны в таблицах 3 и 4. Как показано на дополнительных рисунках 1 и 2, были построены кривые рабочих характеристик приемника (ROC) и диаграммы «Прогнозировано по наблюдаемым данным» для моделей нейронных сетей, оценивающих как только клинические, так и комбинированные клинические и радиомические признаки. В частности, для модели, использующей только клинические данные, площадь под кривой (AUC) для RBNN составила 0,970, а для MLP — 0,785. В комбинированной модели AUC для RBNN достигла 0,992, а для MLP — 0,832. Средняя точность по всем моделям для обучения и тестирования в комбинированном наборе признаков составила $93,31 \pm 4,96$ и $88,17 \pm 2,58$ соответственно. Средняя точность для обучения и тестирования в клиническом наборе признаков составила $91,28 \pm 4,56$ и $87,69 \pm 3,62$ соответственно. Хотя оба набора признаков показали хорошие результаты в задаче прогнозирования в нашей когорте, включение радиомических признаков привело к небольшому увеличению прогностической способности.

Таблица 3. Результаты прогностического моделирования для комбинированных радиомических и клинических характеристик. XGBoost: eXtreme Gradient Boosting; LSVM: Lagrangian Support Vector Machine; Quest: Random Trees, and Quick, Unbiased, Efficient Statistical Tree; MLP-NN: многослойный перцептрон; RBF-NN: нейронная сеть с радиальными базисными функциями.

Алгоритм	Точность обучения (%)	Точность тестирования (%)
Случайные деревья	100.0	89.1
Дерево XGBoost	100.0	91,5
LSVM	89.1	84,5
SVM	90.8	89.3
ЧАЙД	93.7	85.7
MLP-NN	92.0	88.2
RBF-NN	87,5	91.0

Таблица 4. Результаты прогностического моделирования клинических характеристик. XGBoost: eXtreme Gradient Boosting; LSVM: Lagrangian Support Vector Machine; Quest: Random Trees, and Quick, Unbiased, Efficient Statistical Tree; MLP-NN: многослойный перцептрон; RBF-NN: нейронная сеть с радиальными базисными функциями.

Алгоритм	Точность обучения (%)	Точность тестирования (%)
Случайные деревья	95,5	89.7

Дерево XGBoost	100.0	90,5
LSVM	89.6	83.8
SVM	88.1	90.2
ЧАЙД	89.8	85,5
MLP-NN	90.4	82.7
RBF-NN	87,5	91.6

Обсуждение. В настоящем исследовании были объединены радиомические и клинические характеристики межпозвоночного диска для решения задач прогнозирования результатов операций на поясничном отделе позвоночника. Наши результаты показали, что включение радиомических характеристик может улучшить прогнозирование, хотя улучшение в нашем исследовании было незначительным. Хотя дополнительные преимущества в точности прогнозирования, полученные за счет радиомических характеристик, были незначительными в нашем исследовании, это подчеркивает потенциальную ценность включения различных типов данных в клинические прогностические модели. Однако это наблюдаемое преимущество следует рассматривать в контексте его клинической значимости и изменчивости, присущей прогностическому моделированию. Многочисленные предыдущие исследования пытались интегрировать высокопроизводительные методы с многомерными характеристиками для моделирования заболеваний, что привело к многообещающим результатам. Эти характеристики охватывают широкий спектр биологических масштабов, от молекулярного до фенотипического [14]. В то время как радиомические подходы к скелетным мышцам, как правило, сосредоточены на опухолях костей, таких как диагностика заболеваний костей, определение дифференциальной диагностики опухолей, прогнозирование осложнений опухоли и оценка прогноза лечения опухоли на основе патологической классификации [15–18], лишь ограниченное число исследований изучало другие состояния, такие как остеопороз [19], болезнь Альцгеймера [20], остеоартроз височно-нижнечелюстного сустава [21], послеоперационная инфекция и воспаление. Что касается грыжи межпозвоночного диска (ГМД), радиомике до настоящего времени уделялось мало внимания [22]. Одна из известных моделей была разработана ортопедическими хирургами для прогнозирования результатов хирургического лечения ГМД на основе клинических данных. Наше исследование предполагает, что включение радиомических признаков может еще больше улучшить эту модель. Наблюдается всплеск интереса к разработке математических моделей, которые объединяют множество прогностических факторов для прогнозирования результатов лечения пациентов и включают их в компьютеризированные прогностические инструменты. Прогностические модели были широко изучены с точки зрения их разработки, валидации и применения. В первичном звене здравоохранения было разработано множество моделей для прогнозирования исхода болей в спине, но для хирургии позвоночника в третичном звене здравоохранения таких моделей немного [23]. Врумен и др. [24] разработали модель для определения того, будут ли пациенты, первоначально обращающиеся с компрессией нервного корня, в конечном итоге подвергнуты операции на поясничном диске. Недавно в нескольких исследованиях были представлены прогностические модели для прогнозирования клинических

результатов и качества жизни пациентов после операции на шейном отделе позвоночника [25], прогностическая модель для оценки боли и функциональных результатов после операции на поясничном отделе позвоночника [26], а также модель прогнозирования длительности пребывания в стационаре после операции на поясничном отделе позвоночника [27]. Для повышения точности прогнозирования клинических результатов необходимы дополнительные модели для более однородных диагностических групп пациентов, особенно исследования, включающие мультимодальные типы данных. Системы здравоохранения во всем мире генерируют множество источников данных. Несмотря на их сложность, крайне важно установить закономерности и незначительные различия в геномных, радиомических, лабораторных или клинических данных, которые способны надежно различать фенотипы или обеспечивать высокий уровень точности прогнозирования. Данные изображений все чаще обрабатываются с помощью сверточных нейронных сетей (CNN). Используя современные методы искусственного интеллекта, можно объединять мультимодальные типы данных для задач прогнозирования, что позволяет использовать более широкий спектр характеристик пациентов. Такой подход открывает возможность обучения гибридных моделей глубокого обучения с использованием комбинации информации о пациенте из геномных, радиомических и клинических данных. Медицинские работники не полагаются исключительно на один тип данных при принятии решений, и этот подход может помочь обеспечить более полную и точную картину состояния пациента [28]. Что касается радиомических характеристик, мы обнаружили, что характеристики GLCM, статистики первого порядка и группы NGTDM были наиболее прогностическими. Характеристики GLCM и NGTDM представляют собой характеристики более высокого порядка пространственного распределения пиксельных точек по сравнению с 2D и 3D характеристиками формы, что предполагает, что 2D и 3D характеристики первого порядка, видимые невооруженным глазом, недостаточны для адекватного описания изображений грыжи межпозвоночного диска. Вместо этого их необходимо дополнить высокоразмерными характеристиками, которые невозможно различить визуально. Поэтому включение количественных радиомических характеристик, как показано в этом исследовании, может выявить более подробную информацию об изображениях грыжи межпозвоночного диска с различных точек зрения [29]. Следует отметить, что С-реактивный белок (СРБ) еще не установлен в качестве маркера, указывающего на грыжу межпозвоночного диска. Тем не менее, существуют исследования, которые показали, что СРБ связан с послеоперационными результатами при грыже межпозвоночного диска и стенозе позвоночного канала [23, 30, 31]. Поэтому мы включили СРБ в качестве ковариаты в наши прогностические модели. Мы также оценили потенциальные корреляции между радиомическими и клиническими характеристиками. Мы выявили незначительные корреляции между CRP и радиомическими признаками из GLDM, GLRLM и GLZM, которые требуют дальнейшего изучения в будущих исследованиях. Учитывая умеренный характер этих корреляций, отсутствие доказательств причинно-следственных связей и небольшой объем данных, использованный в этом исследовании, мы не акцентировали внимание на этих взаимосвязях в результатах нашего текущего исследования. Следует отметить, что использование радиомических признаков для прогнозирования исходов в хирургии позвоночника ограничено,

что ограничивает нашу возможность сравнивать наши результаты с предыдущими исследованиями. Однако анализ радиомических признаков широко исследовался в последние годы для прогнозирования исходов в исследованиях рака [12, 13]. Дальнейшее изучение прогнозирования исходов на основе радиомики необходимо для подтверждения значимости конкретных радиомических признаков в прогностическом моделировании в исследованиях позвоночника. Важно отметить, что хирургические результаты после операции на поясничном отделе позвоночника определяются не только хирургической процедурой и присущей ей патологией. Послеоперационный уход, включая функциональную реабилитацию, играет важную роль в обеспечении оптимальных результатов. Недавний систематический обзор подчеркнул важность стратегий фокусировки внимания во время реабилитационных упражнений для пациентов с заболеваниями опорно-двигательного аппарата [32]. Исследование показало, что внешняя фокусировка внимания (ВФВ) на эффекте движения более эффективна, чем внутренняя фокусировка внимания (ВФВ) на характеристиках движения, в улучшении выполнения движений, особенно у пациентов с заболеваниями опорно-двигательного аппарата. Это подчеркивает необходимость комплексного послеоперационного ухода, который интегрирует функциональную реабилитацию со стратегиями фокусировки внимания, адаптированными к потребностям каждого отдельного пациента. Хотя наше исследование акцентировало внимание на прогностическом моделировании результатов хирургического вмешательства, будущие исследования должны также изучить влияние таких реабилитационных стратегий на прогностические результаты, предлагая целостный подход к уходу за пациентами. Настоящее исследование имеет несколько ограничений. Во-первых, это ретроспективное одноцентровое исследование с относительно небольшим размером выборки, что требует многоцентровой валидации для обеспечения надежных клинических доказательств. Во-вторых, для извлечения радиомических признаков использовалась только одна последовательность сагиттальных T2WI, в то время как современные исследования показывают, что многопараметрические последовательности МРТ могут предоставить дополнительную информацию о поражениях [33]. В нашем исследовании мы стремились разработать прогностическую модель, которая могла бы применяться ко всем пациентам с грыжей диска, независимо от конкретных подтипов. Такой подход был выбран для обеспечения максимально широкой применимости полученных моделей в клинической практике. Хотя существуют различные подтипы грыжи диска (такие как кальцифицированная/окостеневшая, ограниченная и экструдированная грыжи диска), мы не рассматривали эти подтипы отдельно в нашем анализе. Создание отдельных моделей для каждого подтипа потребовало бы гораздо большего размера выборки и снизило бы осуществимость нашего исследования. Более того, включение различных подтипов в наш анализ лучше отражает реальную клиническую практику, где у пациентов наблюдаются разнообразные проявления грыжи диска. Кроме того, мы не рассматривали показатели исхода, связанные с пациентами, которые количественно оценивают боль или другие характеристики, сообщаемые пациентами, которые могут быть важными клиническими переменными для прогностического моделирования. Кроме того, поскольку у пациентов в разных условиях могут быть существенно разные исходы, обобщаемость инструмента прогнозирования не может быть гарантирована. Одним из

ограничений нашего исследования является ограниченная возможность статистического сравнения точности между двумя наборами моделей. Сравнение производительности моделей машинного обучения, основанных только на клинических переменных и на объединенном наборе данных, может быть подвержено ошибкам I и II типа, что может привести к неточностям в определении наличия или отсутствия значимого эффекта, учитывая небольшой размер выборки значений точности. Хотя наш анализ показал небольшое улучшение прогностической способности при включении радиомических признаков, эти результаты следует интерпретировать с осторожностью. Учитывая наблюдаемые скромные улучшения, для будущих исследований крайне важно тщательно оценить, приводят ли эти минимальные улучшения точности прогнозирования при интеграции радиомических и клинических признаков к значимым клиническим различиям или преимуществам в принятии решений. Это особенно важно в свете существующих проблем. Интерпретация небольших различий в контексте более широкой клинической практики имеет решающее значение. Будущие исследования с использованием больших выборок и проспективных исследований будут иметь решающее значение для подтверждения эффективности сочетания радиомических признаков и клинических переменных в клинических условиях.

Выводы. Хотя наши результаты указывают на заметное улучшение при включении радиомических признаков, общее преимущество в задачах прогнозирования остается неоднозначным. Крайне важно сбалансировать вычислительные затраты и усилия по сбору данных с величиной наблюдаемых преимуществ. Тем не менее, эти результаты могут иметь важное значение для будущих исследований. Целесообразно, чтобы в исследованиях рассматривались многомодальные входные данные для прогностического моделирования, а не полагались исключительно на клинические переменные. Такой подход потенциально может повысить эффективность будущих моделей стратификации клинического риска, основанных в основном на клинических признаках.

Список литературы.

1. Andersson, G. B. J. (1999). Epidemiological features of chronic low-back pain. *The Lancet*, 354(9178), 581–585. [https://doi.org/10.1016/S0140-6736\(99\)01312-4](https://doi.org/10.1016/S0140-6736(99)01312-4)
2. Modic, M. T., & Ross, J. S. (2007). Lumbar degenerative disk disease. *Radiology*, 245(1), 43–61. <https://doi.org/10.1148/radiol.2451051706>
3. Boden, S. D., Davis, D. O., Dina, T. S., Patronas, N. J., & Wiesel, S. W. (1990). Abnormal magnetic-resonance scans of the lumbar spine in asymptomatic subjects. *The Journal of Bone and Joint Surgery*, 72(3), 403–408.
4. Fardon, D. F., Williams, A. L., Dohring, E. J., Murtagh, F. R., Gabriel Rothman, S. L., & Sze, G. K. (2014). Lumbar disc nomenclature. *Spine Journal*, 14(11), 2525–2545. <https://doi.org/10.1016/j.spinee.2014.04.022>
5. Jensen, M. C., Brant-Zawadzki, M. N., Obuchowski, N., Modic, M. T., Malkasian, D., & Ross, J. S. (1994). Magnetic resonance imaging of the lumbar spine in people without back pain. *New England Journal of Medicine*, 331(2), 69–73. <https://doi.org/10.1056/NEJM199407143310201>

6. Weinstein, J. N., Tosteson, T. D., Lurie, J. D., et al. (2006). Surgical vs nonoperative treatment for lumbar disk herniation. *JAMA*, 296(20), 2441–2450. <https://doi.org/10.1001/jama.296.20.2441>
7. Assaf, Y., & Pasternak, O. (2008). Diffusion tensor imaging (DTI)-based white matter mapping. *NeuroImage*, 42(3), 1128–1138.
8. Lambin, P., Rios-Velazquez, E., Leijenaar, R., et al. (2012). Radiomics: Extracting more information from medical images. *European Journal of Cancer*, 48(4), 441–446. <https://doi.org/10.1016/j.ejca.2011.11.036>
9. Gillies, R. J., Kinahan, P. E., & Hricak, H. (2016). Radiomics: Images are more than pictures. *Radiology*, 278(2), 563–577. <https://doi.org/10.1148/radiol.2015151169>
10. Zwanenburg, A., Vallières, M., Abdalah, M. A., et al. (2020). The Image Biomarker Standardisation Initiative. *Radiotherapy and Oncology*, 149, 187–197. <https://doi.org/10.1016/j.radonc.2020.04.016>
11. van Griethuysen, J. J. M., Fedorov, A., Parmar, C., et al. (2017). Computational radiomics system to decode the radiographic phenotype. *Cancer Research*, 77(21), e104–e107. <https://doi.org/10.1158/0008-5472.CAN-17-0339>
12. Aerts, H. J. W. L., Velazquez, E. R., Leijenaar, R. T. H., et al. (2014). Decoding tumour phenotype by noninvasive imaging. *Nature Communications*, 5, 4006. <https://doi.org/10.1038/ncomms5006>
13. Mayerhoefer, M. E., Materka, A., Langs, G., et al. (2020). Introduction to radiomics. *Journal of Nuclear Medicine*, 61(4), 488–495. <https://doi.org/10.2967/jnumed.118.222893>
14. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
15. Schmidhuber, J. (2015). Deep learning in neural networks. *Neural Networks*, 61, 85–117. <https://doi.org/10.1016/j.neunet.2014.09.003>
16. Hastie, T., Tibshirani, R., & Friedman, J. (2017). *The elements of statistical learning* (2nd ed.). Springer.
17. Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.
18. Deo, R. C. (2015). Machine learning in medicine. *Circulation*, 132(20), 1920–1930. <https://doi.org/10.1161/CIRCULATIONAHA.115.001593>
19. Esteva, A., Kuprel, B., Novoa, R. A., et al. (2017). Dermatologist-level classification of skin cancer. *Nature*, 542(7639), 115–118.
20. Choy, G., Khalilzadeh, O., Michalski, M., et al. (2018). Current applications and future impact of machine learning in radiology. *Radiology*, 288(2), 318–328.
21. Park, H. J., Kim, S. S., Lee, S. Y., et al. (2016). Clinical correlation of disc degeneration on MRI. *Spine*, 41(10), E623–E629.
22. Pfirrmann, C. W. A., Metzdorf, A., Zanetti, M., Hodler, J., & Boos, N. (2001). Magnetic resonance classification of lumbar intervertebral disc degeneration. *Spine*, 26(17), 1873–1878.
23. Kjaer, P., Leboeuf-Yde, C., Korsholm, L., & Bendix, T. (2005). Magnetic resonance imaging and low back pain. *Spine*, 30(10), 1173–1180.

24. Turk, O., Kaya, Y., & Arslan, A. (2019). Predicting lumbar disc herniation using machine learning. *Computer Methods and Programs in Biomedicine*, 173, 117–125.
25. Zhang, Y., Chen, X., Wang, L., et al. (2020). Radiomics-based prediction of lumbar disc degeneration. *European Spine Journal*, 29(11), 2700–2708.
26. Kocak, B., Durmaz, E. S., Ates, E., & Kilickesmez, O. (2019). Radiomics in spinal imaging. *Insights into Imaging*, 10, 108.
27. Berenguer, R., Pastor-Juan, M. R., Canales-Vázquez, J., et al. (2018). Radiomics of CT features. *Cancers*, 10(4), 106.
28. Steyerberg, E. W. (2019). *Clinical prediction models*. Springer.
29. Collins, G. S., Reitsma, J. B., Altman, D. G., & Moons, K. G. M. (2015). Transparent reporting of prediction models (TRIPOD). *Annals of Internal Medicine*, 162(1), 55–63.
30. Moons, K. G. M., Altman, D. G., Reitsma, J. B., et al. (2015). Prognosis and prognostic research. *BMJ*, 349, g7393. <https://doi.org/10.1136/bmj.g7393>

Muallif bilan bog‘lanish uchun e-mail	Author's contact email	Email для связи с автором
SaidovKomron56@gmail.com		