

**МЕДИЦИНСКАЯ ИНФОРМАТИКА/MEDICAL INFORMATICS**DOI: <https://doi.org/10.60797/VMED.2026.8.4> EDN: JKBAFA**ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В СППКР ПО ДАННЫМ ЭЛЕКТРОННЫХ МЕДИЦИНСКИХ КАРТ ДЛЯ ФЕНОТИПИРОВАНИЯ И ДИФФЕРЕНЦИАЛЬНОЙ ДИАГНОСТИКИ БАР И БДР**

Обзор

Гоглов И.О.^{1,*}¹ ORCID : 0009-0009-9948-3324;¹ Воронежский государственный университет, Воронеж, Российская Федерация

* Корреспондирующий автор (igoglov[at]inbox.ru)

Аннотация

Введение. Дифференциальная диагностика биполярного аффективного расстройства (БАР; bipolar disorder, BD) и униполярной депрессии (БДР; major depressive disorder, MDD) является сложной задачей, особенно на ранних стадиях заболевания. Ошибки в диагностике приводят к задержке в назначении правильной терапии. Электронные медицинские карты (ЭМК; electronic health records, EHR) представляют собой перспективный источник данных для разработки моделей машинного обучения и систем поддержки принятия клинических решений (СППКР; clinical decision support systems, CDSS) для выявления пациентов с БАР, дифференциальной диагностики БАР и БДР, прогнозирования диагностической конверсии БДР в БАР.

Материалы и методы. Проведен аналитический обзор 12 публикаций, посвящённых фенотипированию БАР с использованием ЭМК, валидации «меток» диагноза, дифференциальной диагностике БАР и БДР, а также прогнозированию диагностической конверсии БДР в БАР. Проанализированы также применения методов NLP и text mining для извлечения симптомов из клинических заметок.

Результаты. Качество «меток» диагноза зависит от строгости фенотипа и может быть улучшено с помощью более детализированных методов, таких как NLP. Прогнозы конверсии БДР в БАР имеют умеренную точность, и значимость точности и качества данных подчеркивается рядом исследований. Модели на основе ЭМК показывают хорошие результаты в задаче дифференциальной диагностики, но требуют качественной записи симптомов в ЭМК.

Выводы. Для успешного применения СППКР в психиатрии необходимо обеспечить качественное фенотипирование, использование более строгих критериев и обязательную внешнюю валидацию фенотипов. Важно также контролировать соблюдение этических норм и конфиденциальности при обработке данных. Системы поддержки принятия клинических решений могут эффективно использоваться для риск-стратификации и триггера в процессе дифференциальной диагностики БАР и БДР.

Ключевые слова: электронные медицинские карты, ЭМК-фенотипирование, биполярное расстройство, униполярная депрессия, машинное обучение, NLP, СППКР.

ARTIFICIAL INTELLIGENCE IN THE CLINICAL DECISION SUPPORT SYSTEMS USING ELECTRONIC CLINICAL RECORDS DATA FOR PHENOTYPING AND DIFFERENTIAL DIAGNOSIS OF BAD AND MDD

Review article

Goglov I.O.^{1,*}¹ ORCID : 0009-0009-9948-3324;¹ Voronezh State University, Voronezh, Russian Federation

* Corresponding author (igoglov[at]inbox.ru)

Abstract

Introduction. The differential diagnosis of bipolar affective disorder (BD) and major depressive disorder (MDD) is a complex task, particularly in the early stages of the illness. Diagnostic errors lead to delays in the prescription of the correct treatment. Electronic medical records (EHR) represent a promising source of data for the development of machine learning models and clinical decision support systems (CDSS) for identifying patients with BPD, differentiating between BPD and MDD, and predicting the diagnostic conversion of MDD to BPD.

Materials and methods. An analytical review was conducted of 12 publications focusing on the phenotyping of BAD using EHR, the validation of diagnostic "labels", the differential diagnosis of BAD and MDD, and the prediction of diagnostic conversion from MDD to BAD. The use of NLP and text mining methods for extracting symptoms from clinical notes was also analysed.

Results. The quality of diagnostic "labels" depends on the specificity of the phenotype and can be improved using more detailed methods, such as NLP. Predictions of conversion from MDD to BAD have moderate accuracy, and the importance of accuracy and data quality is highlighted by a number of studies. Models based on EHR show good results in the task of differential diagnosis, but require high-quality recording of symptoms in the EHR.

Conclusions. For the successful application of CDSS in psychiatry, it is essential to ensure high-quality phenotyping, the use of more strict criteria, and mandatory external validation of phenotypes. It is also important to control compliance with ethical standards and confidentiality when processing data. Clinical decision support systems can be effectively used for risk stratification and as a trigger in the process of differential diagnosis of BAD and MDD.

Keywords: electronic medical records, EMR phenotyping, bipolar disorder, unipolar depression, machine learning, NLP, CDSS.

Введение

Дифференциальная диагностика биполярного аффективного расстройства (bipolar disorder, BD, БАР) и униполярной депрессии (major depressive disorder, MDD, БДР) — важная клиническая задача. На ранних стадиях заболевания депрессивные симптомы могут преобладать, что приводит к ошибочной интерпретации состояния как униполярной депрессии и, как следствие, к задержке в выявлении биполярного спектра. По данным метаанализа, разрыв между началом симптомов и установлением диагноза БАР может составлять несколько лет, что негативно сказывается на своевременности назначения рекомендованной терапии согласно клиническим рекомендациям [1].

Дополнительным риском является то, что у части пациентов с диагнозом униполярной депрессии впоследствии могут возникать маниакальные или гипоманиакальные эпизоды, что приводит к уточнению диагноза БАР. В исследованиях, основанных на данных электронных медицинских карт, антидепрессивная терапия была связана с повышением риска последующего диагноза мании или БАР [2].

В условиях такой сложности и неопределённости диагностики системы поддержки принятия клинических решений становятся важным инструментом. Они могут улучшать процесс дифференциальной диагностики благодаря способности обрабатывать большие объёмы данных и учитывать множество факторов, включая динамику состояния пациента, результаты лабораторных исследований и текстовые записи врачей. Это особенно актуально в контексте биполярного расстройства, где раннее и точное распознавание состояния критично для назначения адекватной терапии.

Одним из перспективных источников данных для СППКР являются ЭМК [3]. В отличие от одноразовых опросников, ЭМК позволяют проводить анализ данных в продольном разрезе и выявлять клинические маркеры, которые могут быть упущены при одном визите, включая сведения, зафиксированные в свободном тексте (с применением технологий обработки естественного языка, NLP).

Цель данного обзора заключается в обобщении методов искусственного интеллекта (ИИ) и машинного обучения (ML), обученных на данных ЭМК, для следующих задач:

1. Фенотипирование БАР и контроль качества «меток».
2. Дифференциальная диагностика БАР и БДР в момент обращения или госпитализации.
3. Прогнозирование диагностической конверсии от БДР к БАР, с акцентом на типы используемых данных, метрики качества, интерпретируемость и переносимость моделей в составе СППКР.

Несмотря на растущий интерес к применению методов машинного обучения и обработки естественного языка в психиатрии, до настоящего времени не было опубликовано значительных обзоров литературы, фокусирующихся на комплексном анализе данной темы. Существующие обзорные исследования частично пересекаются с тематикой, но не охватывают весь спектр задач, рассматриваемых в данной работе. Например, Harvey и соавторы провели скопинг-обзор использования методов NLP для исследований биполярного расстройства, однако их работа сосредоточена на обработке текста и языковых особенностях, а не на построении диагностических моделей на основе ЭМК для задач фенотипирования и дифференциальной диагностики БАР и БДР [11].

Таким образом, актуальность нашего обзора заключается в его сосредоточенности на узкой, но клинически значимой проблеме. Имеющиеся работы либо слишком общие (охватывают различные психические заболевания), либо ограничены отдельными компонентами задачи и не рассматривают весь путь от данных ЭМК до внедрения рекомендательных СППКР.

Терминология

Ниже приведено определение терминологии, которое может быть полезно читателю, не специализирующемуся в области машинного обучения и искусственного интеллекта.

1. СППКР (системы поддержки принятия клинических решений) — информационные системы, которые помогают врачам принимать обоснованные и своевременные клинические решения. Они используют данные о пациентах и алгоритмы машинного обучения для повышения точности диагностики и выбора терапии.

2. ЭМК (электронные медицинские карты) — цифровые версии медицинских карт пациента, которые содержат важные данные о его здоровье: диагнозы, назначения, лабораторные результаты, записи врачей и т. д. ЭМК позволяют анализировать информацию о пациенте в динамике.

3. NLP (обработка естественного языка) — направление в области искусственного интеллекта, позволяющее компьютерам анализировать и интерпретировать текст на естественных языках. В медицине NLP помогает извлекать полезную информацию из свободного текста (например, из заметок врачей о пациентах).

4. Text mining (извлечение информации из текста) — процесс, при котором с помощью алгоритмов извлекается полезная информация из больших объёмов текста; этот подход особенно востребован при обработке медицинских записей.

5. Фенотипирование (в контексте ЭМК) — это процесс формального определения и выделения «случаев» и/или «контролей» заболевания в базе электронных медицинских карт по заранее заданному правилу (алгоритму), которое использует структурированные данные (диагнозы/коды, назначения, госпитализации, лабораторные показатели и др.) и при необходимости неструктурированный текст (NLP), чтобы получить рабочие метки для дальнейшего анализа и обучения моделей.

6. Метрики качества (PPV, NPV, AUC, recall, precision, F-score):

– PPV (positive predictive value) — положительная прогностическая ценность: вероятность того, что положительный результат действительно указывает на наличие заболевания.



– NPV (negative predictive value) — отрицательная прогностическая ценность: вероятность того, что отрицательный результат действительно означает отсутствие заболевания.

– AUC (area under the curve) — площадь под ROC-кривой, отражающая общую эффективность модели.

– recall (полнота) — способность модели выявлять все случаи заболевания.

– precision (точность) — доля корректных положительных предсказаний среди всех положительных предсказаний модели.

– F-score — гармоническое среднее между precision и recall; это делает F-меру сбалансированной метрикой для оценки бинарной классификации.

7. Диагностическая конверсия (diagnostic conversion) — процесс, при котором диагноз пациента меняется со временем (например, от униполярной депрессии (БДР) к биполярному расстройству (БАР)) по мере появления новых симптомов.

8. Логистическая регрессия (logistic regression) — один из базовых методов машинного обучения для бинарной классификации (например, «болен/не болен»). Алгоритм оценивает вероятность принадлежности объекта к классу (например, «пациент с БАР») с помощью логистической функции.

9. Random forest (случайный лес) — ансамблевый метод, использующий множество деревьев решений. Каждое «дерево» обучается на случайной подвыборке данных, а итоговый прогноз формируется агрегированием результатов всех деревьев.

10. SVM (support vector machine) — алгоритм машинного обучения для задач линейной и нелинейной классификации, регрессии и обнаружения аномалий. Если данные не линейно разделимы, SVM использует ядровые функции (kernel), которые переводят данные в пространство большей размерности.

11. XGBoost (extreme gradient boosting) — метод градиентного бустинга, в котором деревья решений строятся последовательно: каждое следующее дерево исправляет ошибки предыдущего. Алгоритм часто демонстрирует высокую производительность на сложных задачах.

12. Многомерная Cox-регрессия (multivariate Cox regression) — статистический метод анализа выживаемости, моделирующий время до наступления события (например, смерти или рецидива заболевания) с учётом нескольких факторов, влияющих на исход.

13. micro-avg и macro-avg — методы усреднения метрик (precision, recall, F-score) в задачах классификации:

– micro-avg: агрегирует предсказания и метки по всем классам и затем вычисляет общую метрику;

– macro-avg: вычисляет метрику отдельно для каждого класса и затем усредняет результаты (все классы имеют одинаковый вес).

14. Методы объяснения моделей (SHAP и break down):

– SHAP — метод объяснения моделей машинного обучения, использующий значения Шепли (теория игр) для оценки вклада каждого признака (например, возраста, лабораторных показателей или симптомов) в предсказание. SHAP повышает интерпретируемость моделей, делая их более прозрачными для клинициста.

– Break down — метод пошагового разложения предсказания модели на вклады признаков. Он помогает понять, какие именно факторы (например, наличие гипомании, возраст или лечение) вносят наибольший вклад в итоговое решение модели.

Логика работы с ЭМК-данными (последовательность этапов)

Чтобы связать между собой все разделы обзора, ниже приведена типичная последовательность построения ИИ-моделей и их внедрения на основе ЭМК:

1. Сбор и подготовка данных ЭМК: структурированные поля (коды, назначения, госпитализации, лабораторные показатели) и, при наличии, тексты клинических заметок.

2. Фенотипирование: разработка и применение ЭМК-фенотипа (правила/алгоритма), по которому формируются группы «случай» (пациенты, которые считаются имеющими целевое состояние) и «контроль» (пациенты, которые не имеют целевого состояния). ЭМК-фенотип может включать комбинацию кодов, назначений, клинического контекста и других признаков.

3. Формирование меток: присвоение каждому пациенту (или эпизоду) классификационной метки на основе выбранного фенотипа (например, БАР vs БДР).

4. Валидация меток: проверка соответствия меток независимому эталону (структурированное интервью/экспертная разметка) и/или конвергентная проверка (например, генетическая согласованность).

5. Обучение и оценка моделей: построение ML-моделей под конкретную клиническую задачу (БАР vs БДР при обращении; прогноз конверсии БДР → БАР) с корректными временными окнами и контролем утечки информации.

6. Оценка клинической применимости: помимо AUC, оценка калибровки, выбор порогов, PPV/NPV при реальной распространённости и анализ клинической пользы.

7. Внедрение в СППКР: использование модели как риск-стратификации/триггера для углублённой диагностики, с мониторингом качества и соблюдением требований безопасности данных.

Качество «меток» диагноза и фенотипирование БАР по данным ЭМК

Качество «меток» диагноза и фенотипирование биполярного аффективного расстройства на основе данных электронных медицинских карт требует особого внимания к специфике психиатрических данных. Формальный диагноз может отражать особенности кодирования и документации, а не клиническую реальность, что делает критически важным этап фенотипирования. Это включает формирование рабочих определений БАР и контрольных групп на основе совокупности данных ЭМК, а не единичного кода. В обзорной работе Smoller отмечается, что основным ограничением использования ЭМК в психиатрии является необходимость валидности фенотипов, поскольку



неполнота симптомов и неоднородность документации требуют алгоритмов, которые комбинируют структурированные данные и клинический текст, с обязательной проверкой полученных определений [3].

На практике фенотипирование БАР по ЭМК осуществляется через несколько стратегий. Во-первых, используются кодовые определения, например, наличие и повторяемость диагностических кодов БАР. Во-вторых, учитываются клинические контексты, такие как назначения препаратов и частота госпитализаций. В-третьих, для повышения точности применяются неструктурированные клинические заметки, из которых извлекаются важные диагностические маркеры с помощью методов NLP и text mining. Это позволяет получать «метки» (например, БАР vs БДР) с контролируемым уровнем достоверности [3].

Исследования показывают, что выбор фенотипа существенно влияет на качество разметки. В работе Castro и соавт. различные варианты фенотипирования БАР были сопоставлены с клиническим интервью (SCID) как «золотым стандартом». Более строгие определения обеспечивали более высокую прогностическую ценность: для NLP-усиленного фенотипа БАР (95-NLP) значение PPV (Positive predictive value) составило 0,85, а для строгого кодового определения (Coded-Strict) — 0,79. Более «широкие» определения демонстрировали меньшую точность, что подчёркивает важность строгих критериев разметки для обучения моделей ИИ [4].

Дополнительную достоверность ЭМК-фенотипов обеспечивает внешняя валидация. Chen и соавт. показали, что автоматизированные фенотипы БАР, полученные из ЭМК, демонстрируют генетическую согласованность с крупными выборками БАР, подтверждая, что корректно построенный фенотип отражает клинически осмысленную структуру расстройства, а не только административное кодирование [5].

Следовательно, при построении систем поддержки принятия клинических решений (СППКР) на основе ЭМК для различения БАР и БДР первоочередной задачей является обеспечение качества исходных «меток». Это включает использование комбинации структурированных данных и клинического текста (NLP), применение строгих фенотипов с ясными правилами включения и исключения, а также проведение валидации фенотипов. Без этих мер качество моделей будет ограничено шумом разметки и смещениями данных ЭМК [3], [4], [5].

Дифференциальная диагностика БАР vs БДР на данных ЭМК в момент обращения/госпитализации

Вторая прикладная задача в рамках ЭМК-подходов — дифференциальная диагностика при обращении/госпитализации, то есть различение депрессивного эпизода при биполярном расстройстве и униполярной депрессии в момент обращения. Важно учитывать, что в ЭМК-моделях часть сигналов может отражать не только «биполярность», но и тяжесть состояния, коморбидность и особенности лечения. Модель должна выступать как «подсказка врачу», помогая заподозрить биполярный спектр у пациента с депрессивной симптоматикой. Качество такого решения зависит от полноты симптомов и клинического контекста в ЭМК, а не только от диагностических кодов.

Исследование Zhu и соавт. на данных госпитального уровня (n=16 311; 2009–2018) сравнивало алгоритмы машинного обучения (логистическая регрессия, random forest, SVM, XGBoost) для различения БАР и БДР. Важной особенностью работы является акцент на интерпретируемости: авторы использовали методы объяснения (SHAP, break down) для демонстрации вклада отдельных признаков, что делает выводы модели более понятными для клинициста.

На тестовой выборке достигалась умеренно высокая дискриминация (AUC = 0,777) при PPV 0,576 и NPV 0,899 [6]. Это типично для клинических задач с неоднородной популяцией: модель лучше «исключает» БАР (высокий NPV), чем подтверждает его (ограниченный PPV).

При исключении самоотчётных симптомов точность модели снижалась (AUC = 0,701 для БАР vs БДР) [6]. Это подчёркивает важность качественной фиксации симптомов в ЭМК: если симптомы документируются неполно или остаются только в свободном тексте, модель, основанная на демографических и рутинных данных, может быть менее эффективной. Для успешного внедрения СППКР критично обеспечить стандартизированную запись симптомов или использовать NLP для извлечения данных из клинических заметок.

В целом результаты показывают, что модели на клинических и ЭМК-данных могут различать БАР и БДР на уровне, потенциально полезном для клинической практики, особенно при наличии симптомных признаков и интерпретируемых факторов. Практическая ценность таких моделей зависит от качества документации, корректности меток и необходимости внешней проверки переносимости на другие учреждения и популяции [3], [6].

Прогноз диагностической конверсии БДР → БАР как «диффдиагноз во времени»

Задача прогноза диагностической конверсии для СППКР на базе ЭМК заключается в выявлении группы повышенного риска среди пациентов с БДР, у которых позднее может быть уточнён диагноз биполярного аффективного расстройства. Важно, чтобы модель прогнозировала вероятность уточнения диагноза, основываясь на данных, доступных на момент диагностики, чтобы избежать «подглядывания в будущее».

В работе Pradier и соавт. рассматривалась клиническая ситуация, когда у пациентов с БДР оценивался риск постановки кода БАР в течение 3 месяцев после начала антидепрессивной терапии. Исследование охватило 67 807 пациентов, и конверсия наблюдалась у 1,36%. Модели логистической регрессии и random forest продемонстрировали среднюю дискриминацию AUC = 0,76 (0,73–0,80), что позволяет выделять пациентов с повышенным риском уточнения диагноза после начала лечения. Такие подходы могут быть полезны для калибровки интенсивности наблюдения и объёма диагностического дообследования в первые месяцы терапии [7].

Сильная сторона методов прогнозирования конверсии заключается в их проверяемости на различных источниках данных. Nestsiarovich и соавт. построили модель 1-годовой конверсии БДР → БАР в формате ОМОР (стандартный формат для хранения и анализа данных наблюдений за состоянием здоровья) и провели масштабную валидацию в международной сети OHDSI. Модель, обученная на пяти базах данных США, показала AUC = 0,689 (0,633–0,745), а на девяти внешних базах AUC = 0,664 (0,570–0,785). В качестве предикторов перехода в БАР выделены: молодой возраст, тяжёлая депрессия, психоз, тревога, злоупотребление ПАВ, мысли о самоповреждении и предшествующие



психические расстройства. Модель позволяла разделять пациентов на группы с различиями риска до 100-кратных, что подчёркивает клиническую полезность даже при умеренных значениях AUC [8].

В исследовании Service и соавт. задача конверсии рассматривалась на горизонте до 5 лет после первичного диагноза БДР, охватывая 13 607 пациентов, из которых 1610 (11,8%) получили диагноз БАР. Использовалась многомерная Соx-регрессия для учета цензурирования и разной длительности наблюдения. Среди предикторов конверсии указывались тяжесть первого эпизода, психоз, госпитализация, семейный анамнез БАР и ассоциации с классами лекарств. Суицидальность, извлечённая из клинических заметок с помощью NLP, также была связана с риском конверсии. Для 5-летнего прогноза AUC = 0,65, с recall 72% и precision 38%, что свидетельствует о способности модели находить значительную долю будущих пациентов с конверсией, но с умеренной долей ложных срабатываний [9].

В целом результаты исследований по конверсии БДР → БАР показывают, что с помощью ЭМК можно строить модели с умеренной дискриминацией, которые полезны не для автоматической постановки диагноза, а для выделения группы повышенного риска и как повод к углублённой диагностике. Качество интерпретации зависит от горизонта прогноза (3 месяца, 1 год, 5 лет), частоты события в популяции и степени внешней валидации на независимых данных [7], [8], [9].

Использование NLP для извлечения симптомов и признаков из клинических заметок ЭМК

Психиатрические данные в электронных медицинских картах (ЭМК) часто неполные для задач ИИ и СППКР, поскольку важная клиническая информация фиксируется в свободном формате, например, описания симптомов и динамика состояния. Это приводит к тому, что один и тот же пациент может выглядеть как «БДР» в структурированных полях, но иметь маркеры биполярного спектра в текстах. Поэтому всё больше исследований применяют методы NLP и text mining для преобразования клинических заметок в формализованные признаки, пригодные для статистики и машинного обучения.

Работа Wu и соавт. показывает, что коды для БДР приемлемы, но точность кодирования для БАР ниже, что увеличивает риск ошибок фенотипирования. Анализ текстов позволяет извлекать депрессивные симптомы с F-score около 0,753–0,774 и выделять пациентов с БДР (recall 0,85, precision 0,69). Это подчёркивает важность NLP для добавления симптомного слоя в ЭМК-аналитику, критичного для дифференциации аффективных расстройств [12].

Daí и соавт. демонстрируют возможность построения моделей на небольших клинических корпусах с использованием современных нейросетевых методов. Они исследовали «скрининг» психиатрических пациентов на выборке 500 пациентов с экспертной разметкой диагнозов. Модели с предобучением превосходят модели «с нуля», показывая выигрыш по micro-avg и macro-avg F-score (на 0,11 и 0,28 соответственно). Это говорит о том, что NLP-признаки из текстов ЭМК могут быть реализованы даже при ограниченных ресурсах, если правильно организована разметка и выбран подход к обучению [10].

Обзор Harvey и соавт. показывает, что из 507 найденных работ в итоговый анализ вошли 35 исследований, сгруппированных по направлениям: предсказание/классификация, языковые маркеры БАР, использование ЭМК для оценки исходов и фенотипирования. Важно, что вопросы этики и конфиденциальности упоминаются не во всех работах (60% исследований), что критично для внедрения NLP в клинику. В целом обзор подтверждает, что анализ языка поддерживает клинические задачи при БАР, но качество, переносимость и этические аспекты остаются ограничениями [11].

В итоге результаты исследований по NLP в ЭМК указывают на три практических вывода для СППКР в задаче БАР vs БДР:

- 1) одних кодов часто недостаточно — текстовый слой ЭМК добавляет клинически значимые симптомы;
- 2) NLP позволяет превращать свободный текст в формализованные признаки, полезные для фенотипирования и предиктивных моделей;
- 3) необходимо контролировать качество разметки, проверять переносимость на новых данных и соблюдать требования конфиденциальности [10], [11], [12].

Заключение

Проведённый обзор показывает, что данные электронных медицинских карт (ЭМК) могут быть использованы для трёх ключевых клинических задач при аффективных расстройствах:

- 1) фенотипирование БАР и контроль качества «меток»;
- 2) поддержка дифференциальной диагностики БАР и БДР в момент обращения/госпитализации;
- 3) риск-стратификация и прогноз диагностической конверсии БДР → БАР. Наиболее устойчивый вывод заключается в том, что качество моделей в значительной степени определяется качеством исходного фенотипа/разметки и полнотой клинических данных, а не только выбором алгоритма.

Основные выводы:

1. Для фенотипирования БАР по ЭМК наиболее перспективны строгие правила включения и многоисточниковые определения (коды + назначения + клинический контекст + текст), а также обязательная валидация меток (клиническая и/или генетическая).

2. В задаче дифференциальной диагностики БАР и БДР модели машинного обучения демонстрируют клинически полезную дискриминацию, однако практическая роль СППКР должна формулироваться как «триггер» для углублённого расспроса и уточнения анамнеза, а не как автоматическая постановка диагноза.

3. Модели прогноза конверсии БДР → БАР дают умеренную точность, но уже позволяют выделять подгруппы повышенного риска; наибольшую ценность имеют подходы с внешней валидацией и проверкой переносимости на независимых наборах данных.



4. NLP является важным дополнением к структурированным полям ЭМК, так как позволяет извлекать симптомные признаки из клинических заметок и повышать информативность признакового пространства моделей.

Ограничения и интерпретация. Результаты ЭМК-исследований требуют осторожной интерпретации из-за неоднородности документации, возможной неполноты симптомов, различий в практиках кодирования и рисков смещения/утечки информации при формировании признаков. Отдельной проблемой является назначения (антидепрессанты, нормотимики, антипсихотики) могут отражать клиническое решение врача и тяжесть состояния, а не быть независимыми «предикторами», что ограничивает причинную интерпретацию моделей. Кроме того, для практического внедрения СППКР одной дискриминации (AUC) недостаточно: необходимы оценка калибровки, выбор клинических порогов, анализ клинической полезности (например, decision-curve/benefit) и мониторинг качества при переносе на новые ЛПУ. Наконец, внедрение должно сопровождаться процедурами защиты данных (де-идентификация, контроль доступа, аудит, оценка риска) и governance-подходом (регламенты, ответственность, обновления модели). Поэтому СППКР целесообразно рассматривать как инструмент поддержки решения, дополняющий клиническое обследование.

Перспективы. Приоритетными направлениями дальнейших работ являются:

1. Стандартизация и воспроизводимость фенотипов (в т.ч. единые правила формирования «меток» и временных окон).
2. Расширение внешней валидации (многоцентровые исследования и переносимость между ЛПУ).
3. Развитие интерпретируемых моделей и отчётности результатов (калибровка, пороги, анализ ошибок).
4. Интеграция NLP в клинические контуры при соблюдении требований конфиденциальности и этики.

Конфликт интересов

Не указан.

Рецензия

Все статьи проходят рецензирование. Но рецензент или автор статьи предпочли не публиковать рецензию к этой статье в открытом доступе. Рецензия может быть предоставлена компетентным органам по запросу.

Conflict of Interest

None declared.

Review

All articles are peer-reviewed. But the reviewer or the author of the article chose not to publish a review of this article in the public domain. The review can be provided to the competent authorities upon request.

Список литературы на английском языке / References in English

1. Scott J. A systematic review and meta-analysis of delayed help-seeking, delayed diagnosis and duration of untreated illness in bipolar disorders / J. Scott, A. Graham, A. Yung [et al.] // *Acta Psychiatrica Scandinavica*. — 2022. — Vol. 146. — № 5. — P. 389–405. — DOI: 10.1111/acps.13490.
2. Patel R. Do antidepressants increase the risk of mania and bipolar disorder in people with depression? A retrospective electronic case register cohort study / R. Patel, P. Reiss, H. Shetty [et al.] // *BMJ Open*. — 2015. — Vol. 5. — № 12. — DOI: 10.1136/bmjopen-2015-008341.
3. Smoller J.W. The use of electronic health records for psychiatric phenotyping and genomics / J.W. Smoller // *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*. — 2018. — Vol. 177. — № 7. — P. 601–612. — DOI: 10.1002/ajmg.b.32548.
4. Castro V.M. Validation of electronic health record phenotyping of bipolar disorder cases and controls / V.M. Castro, J. Minnier, S.N. Murphy [et al.] // *American Journal of Psychiatry*. — 2015. — Vol. 172. — № 4. — P. 363–372. — DOI: 10.1176/appi.ajp.2014.14030423.
5. Chen C.Y. Genetic validation of bipolar disorder identified by automated phenotyping using electronic health records / C.Y. Chen, P.H. Lee, V.M. Castro [et al.] // *Translational Psychiatry*. — 2018. — Vol. 8. — № 1. — P. 86. — DOI: 10.1038/s41398-018-0133-7.
6. Zhu T. Explainable machine-learning algorithms to differentiate bipolar disorder from major depressive disorder using self-reported symptoms, vital signs, and blood-based markers / T. Zhu, X. Liu, J. Wang [et al.] // *Computer Methods and Programs in Biomedicine*. — 2023. — Vol. 240. — DOI: 10.1016/j.cmpb.2023.107723.
7. Pradier M.F. Predicting change in diagnosis from major depression to bipolar disorder after antidepressant initiation / M.F. Pradier, M.C. Hughes Jr, T.H. McCoy [et al.] // *Neuropsychopharmacology*. — 2021. — Vol. 46. — № 2. — P. 455–461. — DOI: 10.1038/s41386-020-00838-x.
8. Nestsiarovich A. Predictors of diagnostic transition from major depressive disorder to bipolar disorder: a retrospective observational network study / A. Nestsiarovich, J.M. Repp, M.E. Matheny [et al.] // *Translational Psychiatry*. — 2021. — Vol. 11. — № 1. — P. 642. — DOI: 10.1038/s41398-021-01760-6.
9. Service S.K. Predicting diagnostic conversion from major depressive disorder to bipolar disorder: an EHR based study from Colombia / S.K. Service, J.F. De La Hoz, A.M. Diaz-Zuluaga [et al.] // *Bipolar Disorders*. — 2025. — Vol. 27. — № 1. — P. 47–56. — DOI: 10.1111/bdi.13512.
10. Dai H.-J. Deep Learning-Based Natural Language Processing for Screening Psychiatric Patients / H.-J. Dai, C.-H. Su, Y.-Q. Lee [et al.] // *Frontiers in Psychiatry*. — 2021. — Vol. 11. — DOI: 10.3389/fpsy.2020.533949.
11. Harvey D. Natural Language Processing Methods and Bipolar Disorder: Scoping Review / D. Harvey, F. Lobban, P. Rayson [et al.] // *JMIR Mental Health*. — 2022. — Vol. 9. — № 4. — DOI: 10.2196/35928.
12. Wu C.-S. Using text mining to extract depressive symptoms and to validate the diagnosis of major depressive disorder from electronic health records / C.-S. Wu, C.-J. Kuo, C.-H. [et al.] // *Journal of Affective Disorders*. — 2020. — Vol. 260. — P. 617–623. — DOI: 10.1016/j.jad.2019.09.044.