

Проект разработки модели полигенных рисков для оценки прогноза развития шизофрении с учетом популяционных особенностей Российской Федерации

Рукавишников Г.В.¹, Скитченко Р.К.^{1,2}, Кибитов А.О.^{1,3}, Пинахина Д.В.¹

¹Национальный медицинский исследовательский центр психиатрии и неврологии им. В. М. Бехтерева, Санкт-Петербург, Россия

²Национальный медицинский исследовательский центр им. В.А. Алмазова, Санкт-Петербург, Россия

³Первый Санкт-Петербургский государственный медицинский университет им. акад. И.П. Павлова, Санкт-Петербург, Россия

Оригинальная статья

Резюме. Понимание генетической архитектуры шизофрении дает представление этиологической гетерогенности заболевания и помогает решить ряд важных вопросов, связанных с ограничениями текущих генетических исследований. Однако, несмотря на данные, базирующиеся на больших когортах, оценка рисков формирования шизофрении на основе генетических данных в настоящее время затруднена в связи с рядом факторов. В значительной степени это связано с этническим разнообразием. Это является особенно актуальным для популяции Российской Федерации, имеющей полиэтнический характер. В результате значительной миграции и активного межэтнического смешения население России представляет собой сложный этнический конгломерат. При этом генетические основы психических заболеваний у полиэтнического и поликультурного населения Российской Федерации остаются недостаточно изученными. Цель статьи: на основании анализа данных о популяционных генетических исследованиях шизофрении предложить проект разработки адаптированной для российской популяции модели с высокими предиктивными показателями. Результаты: Предлагаемый проект направлен на интеграцию в единую базу данных микрочипового генотипирования нескольких когорт пациентов с шизофренией и здоровых добровольцев с последующим изучением в российской популяции полигенных шкал риска шизофрении. В результате реализации проекта впервые в российской популяции будет проведено GWAS с минимально достаточным объемом когорт и выполнено исследование PRS шизофрении, составленных на основании полногеномных исследований ассоциаций в европейских и восточноазиатских когортах и

адаптированных к полиэтнической российской когорте. Также будет разработан и валидизирован мета-предиктор риска шизофрении на основе полигенных шкал и разработана модель оценки риска развития шизофрении, адаптированная к российской популяции.

Ключевые слова: шизофрения, полигенные шкалы риска, полногеномное ассоциативное исследование, популяционные исследования, российская популяция.

Информация об авторах:

Рукавишников Григорий Викторович* - e-mail: grigory_v_r@mail.ru, <https://orcid.org/0000-0002-5282-2036>

Скитченко Ростислав Константинович - e-mail: rost20151995@gmail.com, <https://orcid.org/0000-0002-7203-8206>

Кибитов Александр Олегович - e-mail: druggen@mail.ru, <https://orcid.org/0000-0002-8771-625X>

Пинахина Дарья Владимировна - e-mail: acanthodasha@gmail.com, <https://orcid.org/0000-0001-9896-6556>

Как цитировать: Рукавишников Г.В., Скитченко Р.К., Кибитов А.О., Пинахина Д.В. Проект разработки модели полигенных рисков для оценки прогноза развития шизофрении с учетом популяционных особенностей Российской Федерации...

Конфликт интересов: А.О. Кибитов является заместителем главного редактора.

Финансирование: исследование выполнено при поддержке гранта РФФИ №25-25-00153

Project for the polygenic risk model for the schizophrenia risks prognosis considering the population characteristics of the Russian Federation

Grigory V. Rukavishnikov¹, Rostislav K. Skitchenko^{1,2}, Alexander O. Kibitov^{1,3}, Daria V. Pinakhina¹

¹V.M. Bekhterev National Medical Research Centre for Psychiatry and Neurology.

St. Petersburg, Russia

²National Medical Research Center named after V.A. Almazov, St. Petersburg, Russia

Research article

Abstract. Understanding the genetic architecture of schizophrenia provides insight into the etiologic heterogeneity of the disease and helps to address a number of important issues related to the limitations of current genetic research. However, despite data based on large cohorts, assessing the risks of schizophrenia based on genetic data is currently difficult due to a number of factors. To a significant extent, this is due to ethnic diversity. This is especially relevant for the population of the Russian Federation, which is multiethnic in nature. As a result of significant migration and active interethnic mixing, the population of Russia is a complex ethnic conglomerate. At the same time, the genetic basis of mental illness in the multiethnic and multicultural population of the Russian Federation remains insufficiently studied. Objective: to propose a project for a model with high predictive values adapted to the Russian population based on the analysis of population genetic studies of schizophrenia. Results: The proposed project is aimed at integrating several cohorts of patients with schizophrenia and healthy volunteers into a single database of microarray genotyping, as well as the subsequent study of polygenic risk scores for schizophrenia in the Russian population. As a result of the project, a GWAS with a minimally sufficient cohort size will be performed in the Russian population for the first time, as well as a study of schizophrenia PRS, compiled on the basis of genome-wide association studies in European and East Asian cohorts and adapted to the multiethnic Russian cohort. A meta-predictor of schizophrenia risk based on polygenic scales will also be developed and validated, and a model for assessing the risk of schizophrenia adapted to the Russian population will be developed.

Key words: schizophrenia, polygenic risk scores, genome-wide association study, population studies, Russian population.

Information about the authors:

Grigory V. Rukavishnikov* - e-mail: grigory_v_r@mail.ru, <https://orcid.org/0000-0002-5282-2036>

Rostislav K. Skitchenko - e-mail: rost20151995@gmail.com, <https://orcid.org/0000-0002-7203-8206>

Alexander O. Kibitov - e-mail: druggen@mail.ru, <https://orcid.org/0000-0002-8771-625X>

Daria V. Pinakhina - e-mail: acanthodasha@gmail.com, <https://orcid.org/0000-0001-9896-6556>

To cite this article: Rukavishnikov GV, Skitchenko RK, Kibitov AO, Pinakhina DV. Project for the polygenic risk model for the schizophrenia risks prognosis considering the population characteristics of the Russian Federation...

Conflict of interests: Alexander O. Kibitov - Deputy Editor-in-Chief.

Financing: The research was funded by the Russian Science Foundation grant No. 25-25-00153.

***Автор, ответственный за переписку:** Рукавишников Григорий Викторович - e-mail: grigory_v_r@mail.ru

Corresponding author: Grigory V. Rukavishnikov - e-mail: grigory_v_r@mail.ru

Введение. В настоящее время шизофрения остается одним из наиболее тяжелых, инвалидизирующих психических заболеваний. По данным ВОЗ, от него страдает около 24 млн человек (0,32% населения) в мире [30]. В России, по данным Минздрава за 2021 г., показатель общей заболеваемости шизофренией составил 309 случаев на 100 000 населения (0,309%), а уровень впервые диагностированных случаев — 9,4 на 100 000 в год [4]. Систематический обзор 24-х исследований (2016–2022 гг.) показал, что годовые затраты на одного пациента с шизофренией варьируются от 819 до 94 587 долларов, при этом их большая часть приходится на экономики с высоким доходом [21]. Суммарное экономическое бремя шизофрении в Российской Федерации составляло 196,7 млрд руб. (0,5% валового внутреннего продукта) при общих затратах на лечение около 54 млрд руб. [1]. При этом продолжительность жизни пациентов с шизофренией в среднем на 15-20 лет меньше, чем в среднем по популяции [6]. Связано это не только с суицидами, травматизмом и неблагоприятным образом жизни, но и с уязвимостью общего соматического состояния пациентов и высоким уровнем коморбидности шизофрении с различными соматическими заболеваниями [3,7,16,26]. В связи с этим понимание механизмов развития шизофрении для разработки методов профилактики и терапии является одним из крайне актуальных вопросов в психиатрии и здравоохранении в целом. Механизмы развития шизофрении носят сложный, комплексный характер, включающий значительный вклад генетических и эпигенетических факторов [5,38]. Хорошо известно, что степень семейной предрасположенности к шизофрении значительна, и наследуемость, по полученным на датской популяции данным, составляет около 79% [11]. В результате семейных и

близнецовых исследований была выявлена пропорциональная зависимость риска заболевания от степени генетического родства с человеком, страдающим шизофренией. Так риск приблизительно равен 50% для монозиготных близнецов, 27% - для детей, у которых оба родителя болеют шизофренией, 9% - для родственников первой линии, 2% - для родственников третьей линии [13]. В нескольких крупных исследованиях приемных семей сообщалось о том, что дети, родившиеся у родителей, которые страдали шизофренией, но выросшие в приемных здоровых семьях, имели повышенные показатели заболеваемости по сравнению с контрольной группой усыновленных от здоровых биологических родителей (примерно 10% против 1%) [22,35]. Во всех работах показано, что шизофрения имеет сложную генетическую архитектуру, включающую множество генов и их взаимодействий [12,36]. В связи с полигенным характером заболевания перспективным методом его изучения на основе данных полногеномных ассоциативных исследований (genome wide association studies, GWAS) является вычисление показателей шкал полигенного риска (polygenic risk scores, PRS), определяемых как взвешенная сумма количества аллелей риска для выбранного набора локусов [8]. Однако, несмотря на данные, базирующиеся на больших когортах, оценка рисков формирования шизофрении на основе генетических данных в настоящее время затруднена в связи с рядом факторов. В значительной степени это связано с этническим разнообразием. Разные популяции обладают уникальным генетическим разнообразием, поэтому генетические варианты, ассоциированные с шизофренией, могут различаться между ними [12,15]. При этом большинство GWAS (>80 %) проводятся на индивидах европейского и азиатского происхождения - малые народы и этнические группы (например, народы Сибири, Кавказа, Севера, коренные народы США, Латинской Америки и Африки) в исследованиях почти не представлены. Это приводит к низкой возможности использования полученных на больших когортах PRS и неправильной интерпретации ассоциаций. В группах со смешанным происхождением (например, латиноамериканцы, афро-евразийцы, волжские татары, якуты и др.) интерпретация GWAS и PRS, полученных на других популяционных выборках, также затруднена [10,15]. Так европейская часть PRS может не работать на азиатской части генома, а смешение и гетерозиготность снижают точность LD-моделирования. Понимание генетической архитектуры шизофрении в разных популяциях дает представление о том, существует ли этиологическая гетерогенность заболевания. Указанные проблемы особенно

актуальны для исследований в Российской Федерации. Население РФ составляет около 147 млн человек, весьма неравномерно распределено на большой территории и является полиэтническим (включает более 160 этнических групп восточно- и североευропейского и северо- и среднеазиатского происхождения). Преобладающая этническая группа – русские (~81%) [2]. Все жители Российской Федерации принадлежат к одной из девяти языковых семей: индоевропейской, картвельской, урало-юкагирской, алтайской, эскимосско-алеутской, северокавказской, енисейской, сино-тибетской и чукотско-камчатской. В результате значительной миграции и активного межэтнического смешения население представляет собой сложный этнический конгломерат. При этом генетические основы психических заболеваний у полиэтнического и поликультурного населения Российской Федерации остаются недостаточно изученными. Генетические, геномные и фармакогенетические данные из Российской Федерации мало представлены в международной научной литературе и, следовательно, недоступны для метаанализа и никогда не сравнивались с данными из других популяций.

Гипотеза исследования: комбинация моделей полигенного риска, основанных на различных популяционных выборках, будет иметь более высокую предсказательную способность для населения РФ, чем каждая из моделей в отдельности.

Цель статьи: на основании анализа данных о популяционных генетических исследованиях шизофрении предложить проект разработки адаптированной для российской популяции модели с высокими предиктивными показателями.

Материал и методы. За период с мая по июнь 2025 г. был проведен поиск литературы по базам данных PubMed и Google Scholar по следующим ключевым словам: “schizophrenia”, “genetic studies”, “polygenic risk scores”, “population studies”.

В исследование включались находившиеся в свободном доступе публикации за последние 10 лет, содержащие сведения о популяционных генетических исследованиях шизофрении. Был проведен несистематический обзор выявленных литературных источников с акцентом на вопросы GWAS исследований и применения PRS в изучении шизофрении. Исключались источники, где отсутствовали упоминания генетических исследований шизофрении, также не рассматривались исследования, не содержащие данных об этнических характеристиках выборок.

Основные результаты GWAS исследований шизофрении. Большое исследование GWAS (36 989 случаев и 113 075 контролей) на европейской и восточноазиатской популяциях, опубликованное в 2014 году, выявило 108 локусов в геноме человека, которые были в значительной степени связаны с шизофренией. В соответствии с преобладающими патофизиологическими теориями, среди ассоциированных локусов были гены рецептора дофамина D2 и гены, участвующие в глутаматергической передаче [32]. В работе Lan M. и соавт. (2019) [17] проведен сравнительный анализ генетической архитектуры шизофрении в восточноазиатской и европейской популяциях. Было выявлено 208 значимых локусов, из которых 176 были общими для обеих групп, что подтверждает консервативность биологических механизмов болезни. При этом обнаружен ряд популяционно-специфичных локусов, связанных с нейрональными функциями и иммунным ответом. Исследование по генетике шизофрении, проведенное Tarjinder Singh и соавт. (2022), на многонациональной популяции из 24 248 случаев и 97 322 контролей выявило редкие кодирующие варианты в десяти генах, которые значительно увеличивают риск развития шизофрении. Среди этих генов были идентифицированы *SETD1A*, *CUL1* и *XPO7*, которые до этого не связывались с шизофренией. Эти редкие варианты приводят к 20-52-кратному увеличению риска заболевания. Исследование подчеркивает важность дальнейшего изучения функций этих генов для понимания их роли в патогенезе шизофрении [33]. Другое исследование, проведенное Vassily Trubetskoу и соавт. (2022), сосредоточено на выявлении частых генетических вариаций. Исследователи проанализировали данные от 76 755 человек с шизофренией и 243 649 человек без нее (преимущественно генотипы европейских и восточноазиатских, а также частично латино- и афроамериканских популяционных выборок), выявив 287 участков генома (локусов), связанных с риском развития шизофрении. Среди этих локусов были выявлены 120 генов, потенциально увеличивающих риск заболевания. Большинство этих генов активны только в нейронах и влияют на функции синапсов и организацию синаптических структур [36]. Результаты основных GWAS исследований шизофрении приведены в таблицах 1-2.

Таблица 1. Результаты основных мировых GWAS исследований шизофрении

Table 1. Results of the major worldwide GWAS of schizophrenia

Исследование (год)	Популяция / консорциум	Случаи + Контроль	GWS-локусов	Основные выявленные локусы
Schizophrenia Working Group of the Psychiatric Genomics Consortium (PGC) (2011) [31]	Европейского происхождения	21 856 + 29 839	7	<i>MIR137</i> (1p31), 6p21.3-22.1 (<i>MHC</i>), 18q21.2, <i>CACNA1C</i> , <i>ANK3</i> , <i>ITIH3-ITIH4</i>
Schizophrenia Working Group PGC-2 (2014) [32]	PGC (40 + стран, европейская и восточноазиатская)	36 989 + 113 05	108	<i>DRD2</i> , <i>GRM3</i> , <i>GRIN2A</i> , <i>GRIA1</i> , <i>CACNA1C</i> , <i>CAMKK2</i> , <i>MHC</i> , иммунные гены
SCHEMA (на базе PGC exome) [33]	Многонациональный (европейская (включая отдельные выборки финского и эстонского происхождения), восточно-/южно-	24 248 + 97 322 (экзом)	10 редких PTV-генов	мутированные PTV-варианты: речь о редких, ключ не дан

	азиатская, африканская, ашкенази, африканская, латиноамериканская)			
Schizophrenia Working Group of the PGC; Indonesia Schizophrenia Consortium; Genetic REsearch on schizophreniA neTwork-China and The Netherlands (GREAT-CN) [17]	Восточноазиатская (проведено сравнение с европейской популяцией – данные PGC)	22778 + 35362 (для восточноазиатско й популяции)	19 (для восточноазиат ской популяции)	rs37452893 4 (<i>CACNA2D</i> 2), <i>MIR137</i> , <i>CNTNAP5</i> (для восточноаз иатской популяции)
Schizophrenia Working Group of the PGC-3 [36]	Европейская и восточно-азиатская, частично латино- и афроамериканская	74776 + 101023	287	<i>GRIN2A</i> , <i>SP4</i> , <i>STAG1</i> , <i>FAM120A</i> , <i>BCL11B</i> , <i>CACNA1C</i> , <i>RERE</i> , <i>FOXP1</i> , <i>MYT1L</i> , <i>SLC39A8</i>

Таблица 2. Результаты GWAS исследований шизофрении в российской популяции

Table 2. Results of GWAS of Schizophrenia in the Russian Population

Исследование (год)	Популяция / регион	Случаи + Контроли	Основные выявленные SNP-локусы
Levchenko и соавт. (2021) [19]	Сибирский Федеральный округ (русские)	505 + 503	3p22.2, 8q21.13, 13q14.2
Gareeva и соавт. (2023) [10]	Башкортостан (русские, татары, башкиры)	816 + 989	rs73254185 (4p15.2 – PI4K2B), rs587778384 (20q13.31 – GNAS)

В РФ выполнялись единичные генетические исследования шизофрении, поэтому российская популяция остается не изученной с точки зрения особенностей генетической архитектуры риска заболевания с учетом смешанного характера популяции, ее полиэтничности и мультикультурности при значительном размере. Единственное проведенное исследование [19] в рамках Российского консорциума по психиатрической генетике [9] включало небольшое количество русских пациентов (505 человек) из Сибирского федерального округа и не показало значительных эффектов, связанных с популяционной структурой. Исследование Gareeva A.E. (2023) [10] выполнялось на выборках небольших размеров и при участии индивидов различной этнической принадлежности (русских, татар и башкир, проживающих в Республике Башкортостан). При сравнении основных выявленных локусов на ограниченных выборках можно отметить, что вариант *GNAS* встречается и в российском исследовании (rs587778384), и упоминается в международных данных (хотя *GNAS* не входит в «классические» главные локусы, но упоминается как регуляторный ген, связанный с психическими расстройствами). *CSMD1* часто встречается в международных GWAS и в российских репликациях, но в российских полногеномных исследованиях GWAS его выделяют менее явно. Остальные локусы из российской таблицы (*PI4K2B*, *1q43*) не считаются классическими вариантами для шизофрении на международном уровне, либо ещё недостаточно изучены. Необходимо

отметить, что генетические данные наборов пациентов с шизофренией ФГБНУ НЦПЗ и ИБГ УФИЦ РАН использовались в крупномасштабных GWAS проектах PGC1 и PGC3 соответственно, которые легли в основу ряда полигенных шкал риска для шизофрении.

Тестирование в клинических условиях показывает, что значения полигенного риска, получаемые методом PRS, устойчиво ассоциированы с шизофренией. Zheutlin и соавт. (2019) продемонстрировали 4,6-кратное увеличение риска развития шизофрении у пациентов с наибольшими показателями PRS по сравнению с группой наименьших рисков [40]. В целом, PRS шизофрении на крупных европейский выборках демонстрируют умеренно хорошие прогностические значения ($AUC \approx 0.65\text{--}0.72$, $R^2 \approx 7\text{--}12\%$) [14,39]. Но, несмотря на перспективность PRS для клинической практики, их применение ограничено. Одним из основных препятствий для этого является падение их эффективности при использовании в популяционных выборках, отличающихся по структуре (в том числе, этнической) от выборки GWAS, на основе которого составлялась шкала [25,28]. В поисках решения этой проблемы было показано, что интеграция функциональных данных при разработке шкал увеличивает их кросс-популяционную применимость [20]. Также была предложена разработка мета-предикторов на основе PRS для разнообразных популяций. В частности, этот подход с интеграцией данных полногеномных ассоциаций в 5 популяциях, использовался для разработки улучшенной PRS для ишемической болезни сердца. Ее предиктивная способность оказалась выше во всех исследованных популяциях по сравнению с опубликованными ранее шкалами [27]. Предложены и алгоритмы, основанные на интеграции обоих подходов [37]. Weissbrod и соавт. [37] предлагают в своей работе метод PolyPred, который улучшает эффективность межпопуляционных PRS за счет объединения двух предикторов: нового, который использует функционально обоснованное точное картирование для оценки причинных эффектов (вместо эффектов маркировки), устраняя различия linkage disequilibrium (LD); и ранее опубликованного - BOLT-LMM. При наличии большой обучающей выборки среди неевропейской целевой группы, авторы описывают модель PolyPred+, которая дополнительно включает в себя данные по обучению за пределами Европы. Авторы применили PolyPred к 49 заболеваниям/признакам в 4 популяциях UK Biobank, используя данные обучения UK Biobank British, и получили значимые улучшения по сравнению с BOLT-LMM в диапазоне от +7% у жителей Южной Азии и до +32% у африканцев, что согласуется с моделированием. А при применении

PolyPred+ к 23 заболеваниям/признакам в Британском биобанке Восточной Азии с использованием данных обучения как Британского биобанка Великобритании, так и Биобанка Японии, было получено улучшение на +24% по сравнению с BOLT-LMM и на +12% по сравнению с PolyPred. Аналоги PolyPred и PolyPred+, основанные на сводной статистике, достигли аналогичных улучшений. Еще одним направлением улучшения полигенных шкал является разработка подходов к учету взаимодействия между полиморфизмами в генетической архитектуре заболеваний, которое не учитывается в существующих инструментах. Предварительное тестирование одного из данных подходов на синтетических данных указывает на возможность его использования для выделения новых шкал риска в малых выборках [23]. Эффективность этого подхода для разработки PRS по психиатрическим фенотипам остается недостаточно изученной.

Обсуждение. Ограничения текущих подходов. Как уже упоминалось выше, генетические исследования шизофрении имеют ряд существенных ограничений. Во многом это связано с этиологической гетерогенностью заболевания. Несмотря на единый клинический диагноз, шизофрения представляет собой спектр состояний, которые могут иметь разные патогенетические механизмы. Современные диагностические классификации (DSM-5, МКБ-11) отказались от выделения классических клинических подтипов заболевания, поскольку те не устойчивы диагностически и не имеют чёткой биологической валидности. Тем не менее, исследования указывают на различия в генетической нагрузке при учете клинических характеристик заболевания. Так, например, подтипы, ассоциированные с более выраженной негативной симптоматикой и когнитивным снижением, могут быть более связаны с нарушением нейроразвития [34]. Параноидный подтип часто манифестирует позже и может, по мнению ряда исследователей, иметь меньшую генетическую отягощённость [14,34]. Возраст дебюта также является важным фактором: раннее начало заболевания связано с более высокими показателями PRS и частотой редких структурных вариантов (Copy number variations, CNVs), например, 22q11.2 [29]. Другим важным моментом является то, что крупные генетические исследования шизофрении на настоящее время проводились преимущественно в европейских и восточноазиатских популяционных выборках, что в связи с высокой полигенностью заболевания лишь частично отражает его потенциальные механизмы. При этом генетические основы психических заболеваний у полиэтнического и мультикультурного населения Российской

Федерации остаются недостаточно изученными. Российская Федерация характеризуется высоким популяционно-генетическим разнообразием, включающим множество этнических групп европейского и азиатского происхождения, находящихся в разной степени генетического смешения и имеющих разные уровни ассимиляции. В этом контексте особую значимость приобретает изучение психиатрической генетики в различных регионах: в исторически смешанных европейско-азиатских областях (Волго-Уральский регион, Сибирь) и в более генетически однородных европейских частях страны. Кроме того, генетические, геномные и фармакогенетические данные в отношении психических расстройств в целом и шизофрении, в частности, из Российской Федерации недостаточно представлены в международной научной литературе, недоступны для проведения метаанализов и никогда не сравнивались с данными из других популяций [9]. Для некоторых генетических вариантов гетерогенность, связанная со спецификой популяции, приводит к различной частоте аллелей и неравновесному сцеплению генов (linkage disequilibrium, LD), поэтому ассоциации, выявленные в одной популяции, могут быть не обнаружены в других. Другой проблемой, связанной с популяционными особенностями заболевания, является снижение предиктивной способности PRS.

Возможности перспективных исследований. Полиэтничность российской популяции с учетом существенного влияния мультикультурности представляет собой уникальный объект для изучения методами популяционной генетики, что дает возможность появления новых знаний о генетических системах формирования высокого риска шизофрении. Персонализированный подход к лечению на основе генетических данных может повысить эффективность терапии и снизить частоту нежелательных явлений. Это особенно актуально для фармакогенетических аспектов, где генетические варианты влияют на метаболизм антипсихотических препаратов. Однако реализация этих возможностей в российской популяции требует адаптации существующих полигенных шкал к местным генетическим особенностям. Для этого нами предлагается проект, направленный на интеграцию в единую базу данных микрочипового генотипирования нескольких когорт пациентов с шизофренией и здоровых добровольцев, с последующим изучением в российской популяции PRS шизофрении, составленных по результатам GWAS в европейских и восточноазиатских когортах. На основании полученных результатов планируется разработка адаптированной для российской популяции модели с высокими предиктивными показателями. Для решения

этой задачи в рамках Российского национального консорциума по психиатрической генетике [9] нами реализуется проект по интеграции данных генотипирования трех когорт пациентов с шизофренией и здоровых индивидуумов в российской популяции (г. Москва, г. Томск, г. Уфа). Всего коллекция на текущий момент включает 4717 образцов (1802 случая и 2915 контролей).

В план исследования входит: объединение данных в единую базу данных в пакете hail с контролем их качества и фильтрацией по частоте минорных аллелей. Планируется сравнение частот аллелей вариантов в когортах для оценки возможности их совместного анализа и анализ популяционной структуры выборок, и ее визуализация с помощью метода главных компонент (Principal Component Analysis, PCA).

В дальнейшем будет выполнена валидация полигенных шкал риска шизофрении в российской когорте. Будет создана база данных полигенных рисков шизофрении, составленных по результатам крупномасштабных полногеномных исследований ассоциаций по данным публикаций и международной базы данных полигенных рисков (<https://www.pgscatalog.org/>). Планируется расчет баллов по шкалам для объединенной российской когорты (с помощью PGSC calc (https://github.com/PGScatalog/pgsc_calc) и оценка предсказательной способности шкал (с различными наборами ковариат) с помощью логистической регрессии (пакет statsmodels). В рамках разработки, валидации, и тестирования мета-предиктора риска шизофрении на основе полигенных шкал планируется тренировка предикторов шизофрении с использованием различных алгоритмов машинного обучения (логистическая регрессия, дерево решений, случайный лес, многослойный перцептрон, наивный Байес, k ближайших соседей, метод опорных векторов, AdaBoost и др., пакет sklearn). В качестве признаков будут использоваться значения полигенных рисков, пола, возраста, принципиальных компонент. Валидация предикторов будет проводиться с помощью метода кросс-валидации. Мета-предикторы, прошедшие валидацию, будут протестированы на отдельной под-когорте. Будет проведена оценка связи популяционной принадлежности участников когорты с предсказательной способностью полигенных шкал, разработанных на основе GWAS исследований для восточноазиатских и европейских популяций. Для этого будет проведен модерационный анализ (с использованием модерируемой регрессии), учитывающий взаимодействие вероятности отнесения участника к той или иной популяции по результатам PCA (для этого может

использоваться случайный лес, использующийся в системе `pgsc_calc` для нормализации полигенных рисков [18] с их значениями на основе различных популяций). Также будет проведено сравнение предикторной способности полигенных шкал на основе восточноазиатских и европейских популяций в подвыборках российской когорты с участниками с различной популяционной принадлежностью. В исследовании планируется тестирование инновационного подхода `hiPRS` [23], который подразумевает учет взаимодействия между вариантами и, предположительно, позволяет получить шкалу даже по данным малых когорт. На вход указанный метод использует результаты генотипирования и набор интересующих вариантов при построении модели (для их выбора могут использоваться в том числе данные GWAS). Затем проводится поиск частых множеств (Frequent Itemset Mining, FIM) для построения списка взаимодействий-кандидатов любого заданного порядка с использованием только данных класса риска (пороговая частота задается пользователем). Эти кандидаты ранжируются по взаимной информации с интересующим фенотипом. Алгоритм выбирает k взаимодействий, оптимизируя соотношение этого показателя с мерой избыточности (определяемой сходством взаимодействий), используя подход минимальной избыточности и максимальной релевантности (mRMR) для построения финальной модели на основе логистической регрессии. Оценка предикторной способности существующих полигенных шкал риска будет проводиться как для всей когорты в целом, так и для отдельных под-когорт по центрам исследований. Как упоминалось ранее, Московская и уфимская под-когорты использовались в крупномасштабных GWAS проектах PGC1 и PGC3 соответственно, которые легли в основу ряда полигенных шкал по шизофрении. Возможное наличие обусловленного этим повышения предсказательной способности по шкалам, использующим эти данные, может быть проконтролировано путем тестирования их в под-когорте из Томска, которая не участвовала в данных исследованиях. В случае использования PRS по данным PGC2, возможно также тестирование PRS в когорте из Уфы, которая в PGC2 не участвовала. Тренировка мета-предиктора может вестись на данных когорты из Уфы, которая характеризуется наибольшим популяционным разнообразием. При этом в качестве европейских могут учитываться PRS, использующие данные исследования PGC2. Имеются результаты GWAS в когорте PGC2, из которой исключены представители восточноазиатских популяций [17], подходящие для разработки PRS,

соответствующего европейской популяции для поставленных задач. Перед тренировкой будет случайным образом сделана выборка под-когорты (~0.2 всей когорты) из Уфы для тестирования предиктора, которая не будет использоваться для подбора гиперпараметров и первичной валидации предиктора. Использование этой выборки в тестировании позволит оценить его предсказательную способность в условиях популяционной гетерогенности, характерной для России (когорта из Уфы - наиболее разнообразная из 3 когорт). Подбор гиперпараметров моделей и первичная валидация предсказательной способности может вестись на остальных данных уфимской когорты с использованием кросс-валидации. Далее валидация предиктора может вестись на когорте из Москвы. Окончательное тестирование модели может в данном случае может быть проведено на данных когорты из Томска, которая не участвовала в крупномасштабных GWAS исследованиях, а также на той части когорты из Уфы, которая не использовалась при разработке предиктора и будет более популяционно гетерогенной по сравнению с когортой из Томска. Использование части когорты из Уфы для тестирования возможно, поскольку для предиктора будут использоваться данные PGC2, а не PGC3, в котором когорта из Уфы участвовала. Затем для расширения контрольной группы могут использоваться геномные данные пациентов с биполярным аффективным расстройством (БАР), алкогольной зависимостью и депрессией (ранее созданные в ФГБУ НМИЦ ПН им. В.М. Бехтерева). Они могут быть соединены с когортой из Томска и использованы как отдельные когорты для тестирования предиктора. Использование пациентов с другими психиатрическими расстройствами позволит дополнительно оценить специфичность предиктора к шизофрении и возможность его использования для дифференциации расстройств. Таким образом, по результатам реализации проекта впервые в российской популяции будет проведено GWAS с минимально достаточным объемом когорт, а также выполнено исследование PRS шизофрении, составленных на основании полногеномных исследований ассоциаций в европейских и восточноазиатских когортах и адаптированных к полиэтнической российской когорте. Впервые будет разработан и валидизирован мета-предиктор риска шизофрении на основе полигенных шкал и будет разработана модель оценки риска развития шизофрении, адаптированная к российской популяции.

Заключение. PRS являются перспективным инструментом для изучения шизофрении в клинической практике, однако данный подход обладает рядом существенных ограничений.

Одним из них является падение эффективности при использовании в популяциях, отличающихся по структуре. Особенно актуальным это является для российской популяции в силу ее многообразия и полиэтничности. Нами предлагается проект, направленный на интеграцию в единую базу данных микрочипового генотипирования нескольких когорт пациентов с шизофренией и здоровых добровольцев с последующим изучением в российской популяции полигенных шкал риска шизофрении, составленных по результатам полногеномных исследований ассоциаций в европейских и восточноазиатских когортах. На основании полученных результатов планируется разработка адаптированной для российской популяции модели с высокими предиктивными показателями. Результаты выполнения предлагаемого проекта дадут возможность сравнения данных с результатами изучения других популяций в мире, включая европейскую и восточноазиатскую. Полиэтничность российской популяции с учетом существенного влияния мультикультурности представляет собой уникальный объект для изучения методами популяционной генетики, что даст возможность появлению новых знаний о генетических системах формирования высокого риска шизофрении. При этом генетическая информация, полученная из различных популяций, может привести к разработке более эффективных и персонализированных методов лечения. Например, некоторые генетические варианты могут повлиять на ответ на лекарства, и их знание может помочь в разработке терапий, более подходящих для различных этнических групп.

Благодарности: авторы благодарят Голимбет В.Е., Иванову С.А., Гарееву А.Э., а также коллективы ФГБНУ НЦПЗ, ФГНБУ Томский НИМЦ РАН и ИБГ УФИЦ РАН за возможность организации исследования.

Литература/References:

1. Любов Е.Б., Ястребов В.С., Шевченко Л.С., и соавт. Стоимостный анализ шизофрении в России. Психиатрия и психофармакотерапия им. П.Б. Ганнушкина. 2013;15(01):7-19.

Lyubov EB, Yastrebov VS, Schevchenko LS, et al. The cost analysis of schizophrenia in Russia. Psihiatriya i psihofarmakoterapiya im. P.B. Gannushkina. 2013;15(01):7-19. (In Russ.).

2. Национальный состав населения Российской Федерации, перепись 2020. Федеральная государственная статистическая служба [rosstat.gov.ru]. rosstat.gov; 2025 [проеитировано 25 июня 2025]. Досупно: http://ssl.rosstat.gov.ru/storage/mediabank/Tom5_tab1_VPN-2020.xlsx
National composition of the population of the Russian Federation, 2020 census. Federal State Statistical Service. [rosstat.gov.ru]. rosstat.gov; 2025 [cited 25 June 2025]. Available at: http://ssl.rosstat.gov.ru/storage/mediabank/Tom5_tab1_VPN-2020.xlsx
3. Рукавишников Г.В., Касьянов Е.Д., Жилиева Т.В., Мазо Г.Э. Шизофрения и кардиометаболические нарушения. Журнал неврологии и психиатрии им. С.С. Корсакова. 2021;121(6):132–138.
Rukavishnikov GV, Kasyanov ED, Zhilyaeva TV, Mazo GE. Schizophrenia and cardiometabolic disorders. Zhurnal nevrologii i psikiatrii im. S.S. Korsakova. 2021;121(6):132–138. (In Russ.). <https://doi.org/10.17116/jnevro2021121061132>
4. Шизофрения, клинические рекомендации, возрастная категория: взрослые, 2024. Рубрикатор клинических рекомендаций Министерства здравоохранения РФ [cr.minzdrav.gov.ru]. cr.minzdrav.gov; 2025 [проеитировано 25 июня 2025]. Досупно: https://cr.minzdrav.gov.ru/view-cr/451_3
Schizophrenia, clinical guidelines, age category: adults, 2024. Rubricator of clinical guidelines of the Ministry of Health of the Russian Federation [cr.minzdrav.gov.ru]. cr.minzdrav.gov; 2025 [cited 25 June 2025]. Available at: https://cr.minzdrav.gov.ru/view-cr/451_3
5. Cheng MC, Chien WH, Huang YS, Fang TH, Chen CH. Translational Study of Copy Number Variations in Schizophrenia. Int J Mol Sci. 2021;23(1):457. <https://doi.org/10.3390/ijms23010457>
6. Correll CU, Solmi M, Croatto G, Schneider LK, Rohani-Montez SC, Fairley L, Smith N, Bitter I, Gorwood P, Taipale H, Tiihonen J. Mortality in people with schizophrenia: a systematic review and meta-analysis of relative risk and aggravating or attenuating factors. World Psychiatry. 2022;21(2):248-271. <https://doi.org/10.1002/wps.20994>
7. Dieset I, Andreassen OA, Haukvik UK. Somatic Comorbidity in Schizophrenia: Some Possible Biological Mechanisms Across the Life Span. Schizophr Bull. 2016;42(6):1316-1319. <https://doi.org/10.1093/schbul/sbw028>

8. Escott-Price V. Using polygenic risk score approaches to investigate the common-variant genetic architecture of schizophrenia. V.M. Bekterev review of psychiatry and medical psychology. 2019;(4-1):8-11. <https://doi.org/10.31363/2313-7053-2019-4-1-8-11>
9. Fedorenko OY, Golimbet VE, Ivanova SA, et al. Opening up new horizons for psychiatric genetics in the Russian Federation: moving toward a national consortium. Mol Psychiatry. 2019;24(8):1099-1111. <https://doi.org/10.1038/s41380-019-0354-z>
10. Gareeva AE. Genome-Wide Association Study: Analysis of Association of Polymorphic Loci in 4p15.2 and 20q13.31 Regions with Paranoid Schizophrenia. Russ J Genet. 2023;59:1058–1068. <https://doi.org/10.1134/S1022795423100058>
11. Hilker R, Helenius D, Fagerlund B, Skytthe A, Christensen K, Werge TM, Nordentoft M, Glenthøj B. Heritability of Schizophrenia and Schizophrenia Spectrum Based on the Nationwide Danish Twin Register. Biol Psychiatry. 2018;83(6):492-498. <https://doi.org/10.1016/j.biopsych.2017.08.017>
12. Howrigan DP, Rose SA, Samocha KE, et al. Exome sequencing in schizophrenia-affected parent-offspring trios reveals risk conferred by protein-coding de novo mutations. Nat Neurosci. 2020;23(2):185-193. <https://doi.org/10.1038/s41593-019-0564-3>
13. Iyegbe CO, O'Reilly PF. Genetic origins of schizophrenia find common ground. Nature. 2022;604(7906):433-435. <https://doi.org/10.1038/d41586-022-00773-5>
14. Jonas KG, Lencz T, Li K, Malhotra AK, Perlman G, Fochtmann LJ, Bromet EJ, Kotov R. Schizophrenia polygenic risk score and 20-year course of illness in psychotic disorders. Transl Psychiatry. 2019;9(1):300. <https://doi.org/10.1038/s41398-019-0612-5>
15. Kachuri L, Chatterjee N, Hirbo J, Schaid DJ, Martin I, Kullo IJ, Kenny EE, Pasaniuc B; Polygenic Risk Methods in Diverse Populations (PRIMED) Consortium Methods Working Group; Witte JS, Ge T. Principles and methods for transferring polygenic risk scores across global populations. Nat Rev Genet. 2024;25(1):8-25. <https://doi.org/10.1038/s41576-023-00637-2>
16. Kozumplik O, Uzun S, Jakovljević M. Psychotic disorders and comorbidity: somatic illness vs. side effect. Psychiatr Danub. 2009;21(3):361-7.
17. Lam M, Chen CY, Li Z, et al. Comparative genetic architectures of schizophrenia in East Asian and European populations. Nat Genet. 2019;51(12):1670-1678. <https://doi.org/10.1038/s41588-019-0512-x>

18. Lambert SA, Wingfield B, Gibson JT, et al. The Polygenic Score Catalog: new functionality and tools to enable FAIR research. Preprint. medRxiv. 2024;2024.05.29.24307783. <https://doi.org/10.1101/2024.05.29.24307783>
19. Levchenko A, Kanapin A, Samsonova A, Fedorenko OY, Kornetova EG, Nurgaliev T, Mazo GE, Semke AV, Kibitov AO, Bokhan NA, Gainetdinov RR, Ivanova SA. A genome-wide association study identifies a gene network associated with paranoid schizophrenia and antipsychotics-induced tardive dyskinesia. *Prog Neuropsychopharmacol Biol Psychiatry*. 2021;105:110134. <https://doi.org/10.1016/j.pnpbp.2020.110134>
20. Liang Y, Pividori M, Manichaikul A, et al. Polygenic transcriptome risk scores (PTRS) can improve portability of polygenic risk scores across ancestries. *Genome Biol*. 2022;23(1):23. <https://doi.org/10.1186/s13059-021-02591-w>
21. Lin C, Zhang X, Jin H. The Societal Cost of Schizophrenia: An Updated Systematic Review of Cost-of-Illness Studies. *Pharmacoeconomics*. 2023;41(2):139-153. <https://doi.org/10.1007/s40273-022-01217-8>
22. Lowing PA, Mirsky AF, Pereira R. The inheritance of schizophrenia spectrum disorders: a reanalysis of the Danish adoptee study data. *Am J Psychiatry*. 1983;140(9):1167-71. <https://doi.org/10.1176/ajp.140.9.1167>
23. Massi MC, Franco NR, Manzoni A, et al. Learning high-order interactions for polygenic risk prediction. *PLoS One*. 2023;18(2):e0281618. <https://doi.org/10.1371/journal.pone.0281618>
24. Merikangas AK, Shelly M, Knighton A, Kotler N, Tanenbaum N, Almasy L. What genes are differentially expressed in individuals with schizophrenia? A systematic review. *Mol Psychiatry*. 2022;27(3):1373-1383. <https://doi.org/10.1038/s41380-021-01420-7>
25. Mostafavi H, Harpak A, Agarwal I, Conley D, Pritchard JK, Przeworski M. Variable prediction accuracy of polygenic scores within an ancestry group. *Elife*. 2020;9:e48376. <https://doi.org/10.7554/eLife.48376>
26. Oud MJ, Meyboom-de Jong B. Somatic diseases in patients with schizophrenia in general practice: their prevalence and health care. *BMC Fam Pract*. 2009;10:32. <https://doi.org/10.1186/1471-2296-10-32>

27. Patel AP, Wang M, Ruan Y, et al. A multi-ancestry polygenic risk score improves risk prediction for coronary artery disease. *Nat Med.* 2023;29(7):1793-1803. <https://doi.org/10.1038/s41591-023-02429-x>
28. Ruan Y, Lin YF, Feng YA, et al. Improving polygenic prediction in ancestrally diverse populations. *Nat Genet.* 2022;54(5):573-580. <https://doi.org/10.1038/s41588-022-01054-7>
29. Sada-Fuente E, Aranda S, Papiol S, et al. Correction: Common genetic variants contribute to heritability of age at onset of schizophrenia. *Transl Psychiatry.* 2023;13(1):369. <https://doi.org/10.1038/s41398-023-02651-8>
30. Schizophrenia. WHO official site [who.int]. who; 2025 [Updated 10 January 2022; cited 25 June 2025]. Available at: <https://www.who.int/news-room/fact-sheets/detail/schizophrenia>
31. Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium. Genome-wide association study identifies five new schizophrenia loci. *Nat Genet.* 2011;43(10):969-76. <https://doi.org/10.1038/ng.940>
32. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature.* 2014;511(7510):421-7. <https://doi.org/10.1038/nature13595>
33. Singh T, Poterba T, Curtis D, et al. Rare coding variants in ten genes confer substantial risk for schizophrenia. *Nature.* 2022;604(7906):509-516. <https://doi.org/10.1038/s41586-022-04556-w>
34. Taylor J, de Vries YA, van Loo HM, Kendler KS. Clinical characteristics indexing genetic differences in schizophrenia: a systematic review. *Mol Psychiatry.* 2023;28(2):883-890. <https://doi.org/10.1038/s41380-022-01850-x>
35. Tienari P. Interaction between genetic vulnerability and family environment: the Finnish adoptive family study of schizophrenia. *Acta Psychiatr Scand.* 1991;84(5):460-5. <https://doi.org/10.1111/j.1600-0447.1991.tb03178.x>
36. Trubetskoy V, Pardiñas AF, Qi T, et al. Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature.* 2022;604(7906):502-508. <https://doi.org/10.1038/s41586-022-04434-5>

37. Weissbrod O, Kanai M, Shi H, et al. Leveraging fine-mapping and multipopulation training data to improve cross-population polygenic risk scores. *Nat Genet.* 2022;54(4):450-458. <https://doi.org/10.1038/s41588-022-01036-9>
38. Wockner LF, Morris CP, Noble EP, et al. Brain-specific epigenetic markers of schizophrenia. *Transl Psychiatry.* 2015;5(11):e680. <https://doi.org/10.1038/tp.2015.177>
39. Zhang JP, Robinson D, Yu J, et al. Schizophrenia Polygenic Risk Score as a Predictor of Antipsychotic Efficacy in First-Episode Psychosis. *Am J Psychiatry.* 2019;176(1):21-28. <https://doi.org/10.1176/appi.ajp.2018.17121363>. Erratum in: *Am J Psychiatry.* 2019;176(3):252. <https://doi.org/10.1176/appi.ajp.2019.1763correction>
40. Zheutlin AB, Dennis J, Karlsson Linnér R, et al. Penetrance and Pleiotropy of Polygenic Risk Scores for Schizophrenia in 106,160 Patients Across Four Health Care Systems. *Am J Psychiatry.* 2019;176(10):846-855. <https://doi.org/10.1176/appi.ajp.2019.18091085>

Информация об авторах:

Рукавишников Григорий Викторович — к.м.н., вед. н.с., руководитель отделения социальной нейropsychиатрии ФГБУ «Национальный медицинский исследовательский центр психиатрии и неврологии им. В.М. Бехтерева» (192019, Санкт-Петербург, ул. Бехтерева, д. 3), e-mail: grigory_v_r@mail.ru

Скитченко Ростислав Константинович — к.б.н., м.н.с. отдела грантовых соглашений ФГБУ «Национальный медицинский исследовательский центр психиатрии и неврологии им. В.М. Бехтерева» (192019, Санкт-Петербург, ул. Бехтерева, д. 3), с.н.с. научно-исследовательской лаборатории компьютерного моделирования и искусственного интеллекта ФГБУ «Национальный медицинский исследовательский центр им. В.А. Алмазова» (197341, Санкт-Петербург, ул. Аккуратова, д. 2), e-mail: rost20151995@gmail.com

Кибитов Александр Олегович — д.м.н., гл.н.с., руководитель отделения геномики психических расстройств ФГБУ «Национальный медицинский исследовательский центр психиатрии и неврологии им. В.М. Бехтерева» (192019, Санкт-Петербург, ул. Бехтерева, д. 3), вед. н.с. лаборатории клинической фармакологии аддитивных состояний Института фармакологии им. А.В. Вальдмана ФГБОУ ВО «Санкт-Петербургский государственный» (197022, Санкт-Петербург, ул. Льва Толстого, д. 6-8), e-mail: druggen@mail.ru

Пинахина Дарья Владимировна - к.г.н., научный сотрудник отделения социальной нейropsychиатрии ФГБУ «Национальный медицинский исследовательский центр психиатрии и неврологии им. В.М. Бехтерева» (192019, Санкт-Петербург, ул. Бехтерева, д. 3), e-mail: acanthodasha@gmail.com