

Валидация шкал генетического риска развития артериальной гипертензии на популяции региона Центральной России

Лимонова А.С.¹, Ершова А.И.¹, Киселева А.В.¹, Куценко В.А.^{1,2}, Раменский В.Е.^{1,2}, Вяткин Ю.В.^{1,3}, Сотникова Е.А.¹, Жарикова А.А.^{1,2}, Зайченко М.⁴, Покровская М.С.¹, Шальнова С.А.¹, Мешков А.Н.¹, Драпкина О.М.¹

¹ФГБУ "Национальный медицинский исследовательский центр терапии и профилактической медицины" Минздрава России. Москва; ²ФГБОУ ВО "Московский государственный университет им. М.В. Ломоносова". Москва; ³ФГБОУ ВО "Новосибирский национальный исследовательский государственный университет". Новосибирск; ⁴ФГАОУ ВО "Московский физико-технический институт (национальный исследовательский университет)". Московская область, Долгопрудный, Россия

Цель. Провести валидацию и оценить точность 4 шкал генетического риска (ШГР) артериальной гипертензии (АГ), созданных ранее на европейских выборках, на популяционной выборке населения Ивановской области.

Материал и методы. Для генетического анализа было использовано секвенирование следующего поколения с использованием таргетной панели на выборке региона Центральной России (n=1682) на основе коллекции биобанка. Для валидации отобрано четыре ШГР, ассоциированных с АГ, ранее разработанных для европейской популяции. В качестве метрик качества для моделей регрессии использованы коэффициент детерминации и площадь под ROC-кривой. Проведена дополнительная валидация с включением всех вариантов нуклеотидной последовательности, независимо от уровня их неравновесного сцепления. Составлена объединенная ШГР на основе коэффициентов из отдельных ШГР методом C+T (clumping + thresholding).

Результаты. В рамках исследования было продемонстрировано, что предсказательная способность ранее разработанных ШГР при использовании для популяции региона Центральной России ниже, чем в оригинальных исследованиях, доля объясненной дисперсии составила 0,5-0,8%. Наилучшую предсказательную способность (доля объясненной дисперсии — 2,5%) продемонстрировало применение ранее разработанной ШГР (Evangelou E, et al., 2018), включающей наибольшее количество вариантов нуклеотидной последовательности (n=852).

Заключение. ШГР АГ, разработанные на европейских выборках, не рекомендуется использовать у представителей российской по-

пуляции без предварительной валидации. Для создания собственных ШГР не рекомендуется объединять статистические параметры (β -коэффициенты и p-value) из различных ШГР с целью создания собственных ШГР.

Ключевые слова: артериальная гипертензия, шкала генетического риска, популяционное исследование, валидация.

Отношения и деятельность. Научно-исследовательская работа "Разработка интегрированных систем прогнозирования в персонализированной медицине". Рег. № 121021700364-1.

Поступила 18/10-2023

Рецензия получена 21/10-2023

Принята к публикации 08/11-2023



Для цитирования: Лимонова А.С., Ершова А.И., Киселева А.В., Куценко В.А., Раменский В.Е., Вяткин Ю.В., Сотникова Е.А., Жарикова А.А., Зайченко М., Покровская М.С., Шальнова С.А., Мешков А.Н., Драпкина О.М. Валидация шкал генетического риска развития артериальной гипертензии на популяции региона Центральной России. *Кардиоваскулярная терапия и профилактика*. 2023;22(12):3801. doi:10.15829/1728-8800-2023-3801. EDN ННХАРО

*Автор, ответственный за переписку (Corresponding author):

e-mail: limonova-alena@yandex.ru

[Лимонова А.С.* — н.с. лаборатории клиномики, ORCID: 0000-0003-1500-3696, Ершова А.И. — д.м.н., лаборатория клиномики, руководитель, зам. директора по фундаментальной науке, ORCID: 0000-0001-7989-0760, Киселева А.В. — к.б.н., в.н.с., Институт персонализированной терапии и профилактики, лаборатория молекулярной генетики, руководитель, ORCID: 0000-0003-4765-8021, Куценко В.А. — с.н.с., лаборатория биostatистики отдела эпидемиологии хронических неинфекционных заболеваний, аспирант механико-математического факультета, ORCID: 0000-0001-9844-3122, Раменский В.Е. — в.н.с., руководитель лаборатории геномной и медицинской биоинформатики, Институт персонализированной терапии и профилактики, доцент факультета биоинженерии и биоинформатики, ORCID: 0000-0001-7867-9509, Вяткин Ю.В. — программист, Институт персонализированной терапии и профилактики, лаборатория геномной и медицинской биоинформатики, старший преподаватель факультета естественных наук, кафедра молекулярной биологии и биотехнологии, ORCID: 0000-0002-9056-8796, Сотникова Е.А. — с.н.с. лаборатории молекулярной генетики, Институт персонализированной терапии и профилактики, ORCID: 0000-0002-8395-4146, Жарикова А.А. — к.б.н., лаборатория молекулярной генетики, Институт персонализированной терапии и профилактики, в.н.с., старший преподаватель факультета биоинженерии и биоинформатики, ORCID: 0000-0003-0723-0493, Зайченко М. — аспирант, Физтех-школа биологической и медицинской физики, ORCID: 0000-0002-2798-9811, Покровская М.С. — к.б.н., в.н.с., руководитель Лаборатории "Банк биологического материала", Институт персонализированной терапии и профилактики, ORCID: 0000-0001-6985-7131, Шальнова С.А. — д.м.н., профессор, руководитель отдела эпидемиологии хронических неинфекционных заболеваний, ORCID: 0000-0003-2087-6483, Мешков А.Н. — д.м.н., руководитель Института персонализированной терапии и профилактики, ORCID: 0000-0001-5989-6233, Драпкина О.М. — д.м.н., профессор, академик РАН, директор, ORCID: 0000-0002-4453-8430].

Validation of genetic risk scores for hypertension in the Central Russian population

Limonova A. S.¹, Ershova A. I.¹, Kiseleva A. V.¹, Kutsenko V. A.^{1,2}, Ramensky V. E.^{1,2}, Vyatkin Yu. V.^{1,3}, Sotnikova E. A.¹, Zharikova A. A.^{1,2}, Zaichenoka M.⁴, Pokrovskaya M. S.¹, Shalnova S. A.¹, Meshkov A. N.¹, Drapkina O. M.¹

¹National Medical Research Center for Therapy and Preventive Medicine. Moscow; ²Lomonosov Moscow State University. Moscow; ³Novosibirsk National Research State University. Novosibirsk; ⁴Moscow Institute of Physics and Technology (National Research University). Moscow region, Dolgoprudny, Russia

Aim. To validate and evaluate the accuracy of 4 genetic risk scores (GRSs) for hypertension (HTN), previously created on European samples, on a population sample of the Ivanovo Oblast.

Material and methods. For genetic analysis, targeted next-generation sequencing was used on a sample of the Central Russia (n=1682) based on the biobank collection. Four GRSs associated with HTN, previously developed for the European population, were selected for validation. The coefficient of determination and the area under the ROC curve were used as quality metrics for regression models. Additional validation was carried out to include all nucleotide sequence variants, regardless of linkage disequilibrium level. A combined GRS was compiled based on coefficients from individual GRSs using the clumping + thresholding (C+T) method.

Results. The study demonstrated that the predictive value of previously developed GRSs when used for Central Russian population is lower than in the original studies. The proportion of explained variance was 0,5-0,8%. The best predictive ability (proportion of explained variance — 2,5%) was demonstrated using previously developed GRSs (Evangelou E, et al., 2018), which includes the largest number of nucleotide sequence variants (n=852).

Conclusion. GRSs for HTN, developed on European samples, is not recommended for Russian population without preliminary validation. To create original GRSs, combining statistical parameters (β -coefficients and p-value) from different GRS is not recommended.

Keywords: hypertension, genetic risk score, population study, validation.

Relationships and Activities. Research work "Development of integrated forecasting systems in personalized medicine". Reg. No. 121021700364-1.

Limonova A. S.* ORCID: 0000-0003-1500-3696, Ershova A. I. ORCID: 0000-0001-7989-0760, Kiseleva A. V. ORCID: 0000-0003-4765-8021, Kutsenko V. A. ORCID: 0000-0001-9844-3122, Ramensky V. E. ORCID: 0000-0001-7867-9509, Vyatkin Yu. V. ORCID: 0000-0002-9056-8796, Sotnikova E. A. ORCID: 0000-0002-8395-4146, Zharikova A. A. ORCID: 0000-0003-0723-0493, Zaichenoka M. ORCID: 0000-0002-2798-9811, Pokrovskaya M. S. ORCID: 0000-0001-6985-7131, Shalnova S. A. ORCID: 0000-0003-2087-6483, Meshkov A. N. ORCID: 0000-0001-5989-6233, Drapkina O. M. ORCID: 0000-0002-4453-8430.

*Corresponding author: limonova-alena@yandex.ru

Received: 18/10-2023

Revision Received: 21/10-2023

Accepted: 08/11-2023

For citation: Limonova A. S., Ershova A. I., Kiseleva A. V., Kutsenko V. A., Ramensky V. E., Vyatkin Yu. V., Sotnikova E. A., Zharikova A. A., Zaichenoka M., Pokrovskaya M. S., Shalnova S. A., Meshkov A. N., Drapkina O. M. Validation of genetic risk scores for hypertension in the Central Russian population. *Cardiovascular Therapy and Prevention*. 2023;22(12):3801. doi:10.15829/1728-8800-2023-3801. EDN HNXAPO

АГ — артериальная гипертензия, АД — артериальное давление, ВВП — вариант(-ы) нуклеотидной последовательности, ДИ — доверительный интервал, САД — систолическое АД, ШГР — шкала генетического риска, ЭССЕ-РФ — Эпидемиология сердечно-сосудистых заболеваний и их факторов риска в регионах Российской Федерации, AUC — area under the curve (площадь под ROC-кривой), R² — коэффициент детерминации, C+T — clumping + thresholding, GWAS — genome-wide association studies.

Ключевые моменты

Что известно о предмете исследования?

- Применение шкал генетического риска (ШГР) требует предварительной валидации.

Что добавляют результаты исследования?

- Предсказательная способность ранее разработанных на европейской популяции ШГР артериальной гипертензии при использовании на представителях популяции региона Центральной России ниже, чем в оригинальных исследованиях.
- При разработке собственных ШГР рекомендуется включение большего количества вариантов нуклеотидной последовательности и применение данных полногеномного поиска ассоциаций, в т.ч. на представителях российской популяции.

Key messages

What is already known about the subject?

- The use of genetic risk scores (GRS) requires preliminary validation.

What might this study add?

- The predictive value of GRSs for hypertension previously developed in the European population when used on Central Russian population is lower than in the original studies.
- When developing original GRSs, larger number of nucleotide sequence variants and data from a genome-wide association studies, including on representatives of the Russian population, should be used.

Введение

Артериальной гипертонией (АГ) страдают 1,28 млрд человек во всем мире¹. Она является одним из ведущих, но при этом модифицируемым, факторов риска развития сердечно-сосудистых заболеваний (включая ишемическую болезнь сердца, инфаркт миокарда, инсульт, сердечную недостаточность, заболевания периферических артерий и фибрилляцию предсердий), а также поражения почек, глаз и когнитивных нарушений [1]. АГ имеет комплексную этиологию, включающую генетическую предрасположенность и факторы окружающей среды [2]. С помощью исследований полногеномного поиска ассоциаций GWAS (genome-wide association studies) выявлено множество локусов и вариантов, как частых, так и редких, связанных с уровнем артериального давления (АД) [3, 4]. Понимание генетической предрасположенности к гипертонии могло бы стать полезным инструментом для выявления среди молодых лиц группы повышенного риска развития АГ с целью реализации для нее особых профилактических мероприятий и наблюдений для своевременной установки диагноза. Согласно данным Всемирной организации здравоохранения, лишь 42% случаев заболевания АГ оказываются диагностированными¹, что подчеркивает важность обсуждаемой стратегии для своевременного выявления заболевания.

Однако каждый в отдельности генетический вариант имеет слабое влияние на развитие многофакторных заболеваний, объясняет очень малую долю наследственности в развитии заболевания и, таким образом, имеет ограниченную предсказательную ценность. Одним из решений повышения предсказательной способности генетического тестирования является объединение информации о нескольких вариантах нуклеотидной последовательности (ВНП) в единую систему оценки риска, часто называемую шкалой генетического риска (ШГР) [5]. С помощью ШГР, включавшей >1 млн ВНП, Vaura F, et al. (2021) продемонстрировали увеличение риска развития не только АГ, но и сердечно-сосудистых заболеваний [6]. Более того, было показано увеличение предсказательной способности модели для оценки риска развития АГ при добавлении к клинической информации данных ШГР. Тем не менее стоит отметить, что интерпретация результатов ШГР по сравнению с генетической диагностикой моногенной патологии не является однозначной и позволяет определить лишь вероятность развития заболевания, а не его точный прогноз. На данный момент клинические рекомендации по применению данного подхода отсутствуют. Как заключают авторы недавно

опубликованного рекомендательного документа Американского колледжа медицинской генетики и геномики (American College of Medical Genetics and Genomics), для внедрения ШГР в клиническую практику необходимы дальнейшие проспективные исследования [7].

Применение ШГР на этнической группе, отличной от группы, на которой разработана ШГР, может приводить к неверным оценкам [8], что отчасти обусловлено различиями в неравновесном сцеплении ВНП (linkage disequilibrium) и частотами аллелей [9, 10]. Целью настоящего исследования было оценить воспроизводимость ШГР АГ, ранее разработанных на европейских популяциях, у лиц, проживающих на территории Центральной России.

Материал и методы

Выборка. В настоящее исследование была включена репрезентативная выборка населения Ивановской области из исследования ЭССЕ-РФ (Иваново) (Эпидемиология сердечно-сосудистых заболеваний и их факторов риска в регионах Российской Федерации) [11]. После исключения из анализа родственников в исследование включено 1682 человека (медиана возраста и интерквартильный размах составили, соответственно, 49 [39; 56] лет). Из них 629 мужчин 44 [33; 53] лет и 1053 женщины 52 [43; 57] лет.

В рамках исследования ЭССЕ-РФ диагноз АГ устанавливали при положительном ответе на вопрос "Принимали ли Вы препараты, снижающие АД в течение последних 2 нед.?" и/или при уровне систолического АД (САД) ≥ 140 мм рт.ст. и/или диастолического АД ≥ 90 мм рт.ст.

Секвенирование. Геномную дезоксирибонуклеиновую кислоту (ДНК) выделяли из образцов цельной крови с использованием набора QIAamp DNA Blood Mini Kit (Qiagen, Германия). Для измерения концентрации ДНК использовали флуориметр Qubit 4.0 (Thermo Fisher Scientific, США). Молекулярно-генетический анализ проводили с помощью диагностической панели для прогнозирования хронических неинфекционных заболеваний (ишемической болезни сердца, АГ, острого нарушения мозгового кровообращения), разработанной в ФГБУ "НМИЦ ТПМ" Минздрава России для анализа с помощью секвенирования следующего поколения на приборе NextSeq 550 (Illumina, США). Все этапы были выполнены в соответствии с протоколами производителей.

Выбор ШГР. На основании анализа литературных данных были отобраны 4 ШГР АГ [4, 12–14], разработанные на европейской популяции и опубликованные до начала проведения исследования (в 2019г). Критериями выбора ШГР являлись: количество включенных ВНП (не >1 тыс.), размер выборки в деривационном исследовании не <100 тыс. человек. Обобщенная информация по данным ШГР представлена в таблице 1.

Биоинформатический анализ. Чтения с парными концами в формате fastq были выровнены на референсный геном GRCh37/hg19. Обработка данных и оценка контроля качества выполнялись с помощью специально разработанного пайплайна [15] на базе GATK v.3.8 [16]. ШГР вычислены для каждого образца посредством суммирования эффектов каждого ВНП с применением β -коэффициентов и значений p-value из оригинальных исследований.

¹ World Health Organization. Hypertension. <https://www.who.int/news-room/fact-sheets/detail/hypertension> (17 September 2023).

Таблица 1

Параметры отобранных для данного исследования ШГР АГ

Источник (автор публикации, год)	Размер выборки, чел.	Количество ВНП, включенных в ШГР в оригинальном исследовании	Количество ВНП в настоящем исследовании	Метрики предсказательной способности ШГР, представленные в статье	R ² в настоящем исследовании, %
Ehret GB, et al. (2011) [12]	200000	29	28	R ² = 0,9%	0,7
Ehret GB, et al. (2016) [13]	201529	66	54	R ² = 3,4	0,5
Warren HR, et al. (2017) [14]	140668	152	148	для верхнего квинтиля шанс наличия АГ повышен в >2 раза, по сравнению с нижним (ОШ 2,32, 95% ДИ: 1,76-3,06)	0,8
Evangelou E, et al. (2018) [4]	757601	901	255	R ² = 11,2%	0,8

Примечание: АГ — артериальная гипертензия, ДИ — доверительный интервал, ВНП — вариант нуклеотидной последовательности, ОШ — отношение шансов, ШГР — шкала генетического риска.

Таблица 2

Анализ площади под кривой (AUC) для наличия АГ
и распространенность АГ для нижнего и верхнего децилей ШГР

Источник (автор публикации, год)	Значение AUC, % р значение	Распространенность АГ для нижнего дециля	Распространенность АГ для десятого дециля	р-значение для распространенности АГ в нижнем и верхнем децилях
Ehret GB, et al. (2011) [12]	53,15; (50,38-53,15) 0,026	50,89 (43,10-58,65)	58,58 (50,76-66,09)	0,190
Ehret GB, et al. (2016) [13]	52,61; (49,84-52,61), 0,046	47,93 (40,20-55,73)	56,21 (48,38-63,82)	0,157
Warren HR, et al. (2017) [14]	52,89; (50,12-52,89) 0,041	51,48 (43,68-59,23)	60,95 (53,16-68,35)	0,100
Evangelou E, et al. (2018) [4]	54,63; (51,88-54,63) 0,001	44,97 (37,32-52,80)	65,68 (58,00-72,80)	<0,001

Примечание: АГ — артериальная гипертензия, ШГР — шкала генетического риска, AUC — area under the curve (площадь под ROC-кривой).

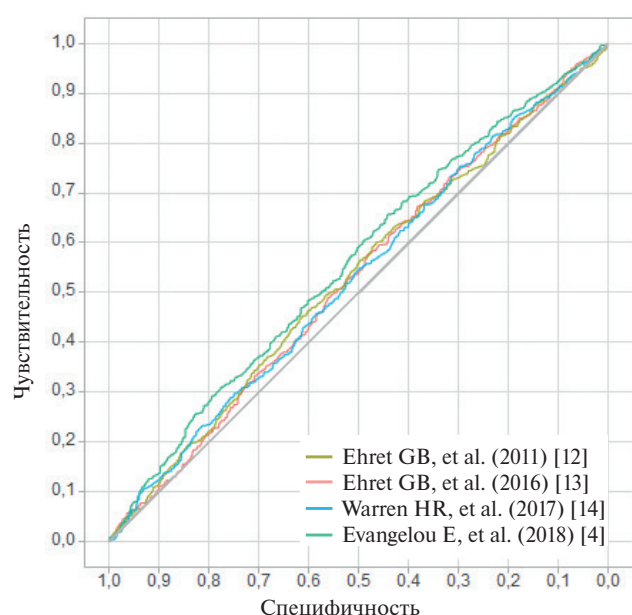


Рис. 1 ROC-кривая для четырех ШГР.

Примечание: ШГР — шкала генетического риска, AUC — area under the curve (площадь под ROC-кривой). Цветное изображение доступно в электронной версии журнала.

Статистический анализ. Для статистического анализа были использованы инструменты языка R v. 4.1. Анализ ассоциаций ШГР с АГ проведен при помощи логистической регрессии с поправкой на пол, возраст и прием антигипертензивной терапии. В качестве метрик качества для моделей регрессии использованы частичный коэффициент детерминации (R²) и площадь под ROC-кривой (area under curve — AUC). Уровень значимости принят равным 0,05.

На следующем этапе анализа была составлена объединенная ШГР на основе коэффициентов из отдельных ШГР методом C+T (clumping + thresholding) [5]. На этапе "clumping" были оставлены только ВНП, имевшие максимальную парную корреляцию с остальными ВНП не >0,1. Оставшиеся ВНП на этапе "thresholding" были подвергнуты пошаговому включению в ШГР на основе медиан р-значений из рассмотренных ШГР. В качестве β-коэффициента использовался медианный β-коэффициент из рассматриваемых ВНП (в том случае, когда один и тот же ВНП использован в нескольких ШГР). Ассоциация ШГР и АГ оценивалась на каждом шаге.

На заключительном этапе проведена дополнительная валидация ШГР Evangelou E, et al. (2018) [4]. Были ослаблены предположения на корреляционную структуру данных, наложенные ранее: включены все ВНП неза-

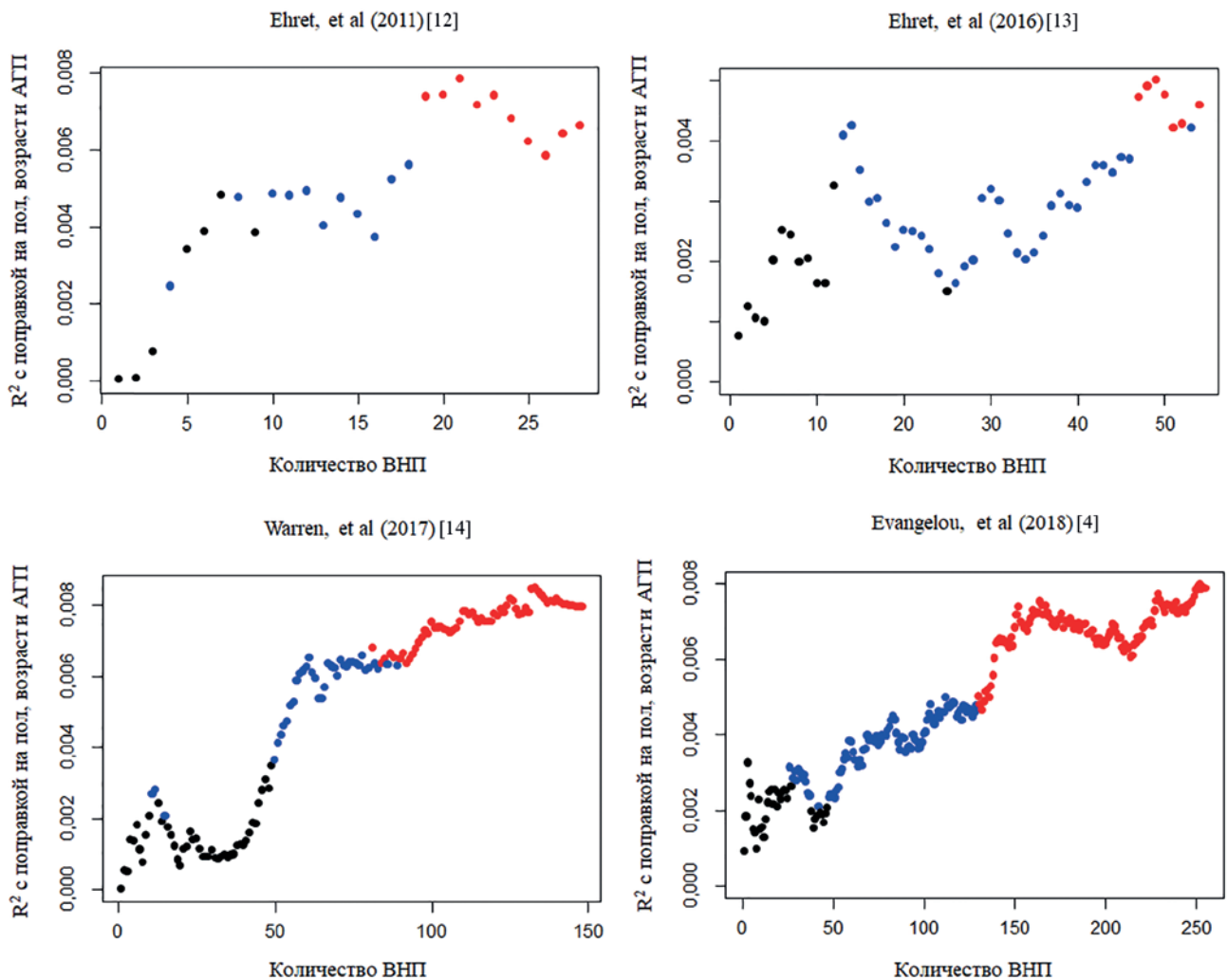


Рис. 2 Результаты линейной регрессии для ассоциации уровня САД со значениями ШГР.

Примечание: для корреляции Пирсона между САД и значением ШГР (из исследований [4, 12-14]: черный — $p > 0,05$, синий — $p < 0,05$, красный — $p < 0,005$. АГП — прием антигипертензивных препаратов, ВВП — вариант нуклеотидной последовательности, САД — систолическое артериальное давление, ШГР — шкала генетического риска, R^2 — коэффициент детерминации. Цветное изображение доступно в электронной версии журнала.

висимо от уровня их неравновесного сцепления. Кроме того, в ШГР включены ВВП, которые ассоциированы с диастолическим и пульсовым давлением.

Результаты

Распространенность АГ в выборке составила 53,8% (902 чел.), из них 64,2% (572 чел.) принимали гипотензивную терапию. Средние показатели САД составили 136 ± 19 мм рт.ст., диастолического АД — 84 ± 10 мм рт.ст.

Первым этапом был проведен анализ силы ассоциации между исследуемыми ШГР и АГ с применением AUC (рисунок 1, таблица 2), AUC варьировала от 52,61% до 54,63%. При этом статистически значимая распространенность АГ для нижнего и верхнего децилей ШГР была установлена только для ШГР из исследования Evangelou E, et al. (2018) [4]. Результаты регрессионного анализа, представленные в виде коэффициента детерминации R^2

с поправкой на пол, возраст и прием антигипертензивной терапии, представлены на рисунке 2. Значение R^2 для всех 4-х ШГР представлено в таблице 1.

Вторым этапом анализа методом С+Т была составлена объединенная ШГР, включавшая 404 ВВП, что позволило достичь увеличения R^2 до 1,5% (рисунок 3). Шанс иметь АГ достоверно увеличивался в среднем в 1,020 раза (95% доверительный интервал (ДИ): 1,012-1,028, $p < 0,001$) на каждый балл ШГР. Разница в распространенности АГ между нижним и верхним децилем составила 17,8% (95% ДИ: 6,7-28,8, $p = 0,004$) (рисунок 4). AUC составила 56,01% (95% ДИ: 53,27-58,76%).

На заключительном этапе проведена дополнительная валидация ШГР Evangelou E, et al. (2018). Всего было включено 852 ВВП (для 49 ВВП отсутствовали данные секвенирования). Полученная ШГР была значимо ($p < 0,001$) ассоциирована с увеличением САД. Полученная ШГР объясняет 2,5%

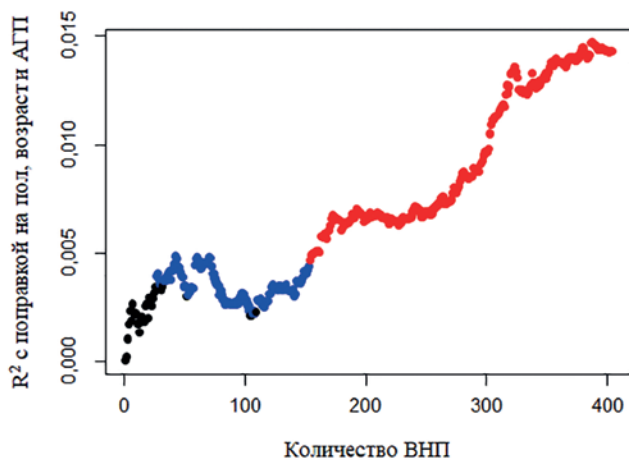


Рис. 3 Результаты линейной регрессии для ассоциации уровня САД со значением объединенной ШГР на основе 404 ВВП. Примечание: для корреляции Пирсона между САД и значением ШГР: черный — $p > 0,05$, синий — $p < 0,05$, красный — $p < 0,005$. АГП — прием антигипертензивных препаратов, ВВП — вариант нуклеотидной последовательности, САД — систолическое артериальное давление, ШГР — шкала генетического риска, R^2 — коэффициент детерминации. Цветное изображение доступно в электронной версии журнала.

дисперсии САД (рисунок 5), AUC составила 56,01% (95% ДИ: 53,27-56,01%). Одно стандартное отклонение ШГР соответствовало увеличению САД на 1,3 мм рт.ст. Распространенность АГ для нижнего дециля составила 46,75% (95% ДИ: 39,04-54,56%), для верхнего — 64,50 % (95% ДИ: 56,78-71,69%).

Обсуждение

В настоящей работе впервые проведена валидация ШГР АГ, разработанных на европейских выборках, у представителей российской популяции. Доля объясненной дисперсии для ШГР, включавшей наименьшее количество ВВП [12], сопоставима с результатами оригинальной статьи. Однако полученные нами результаты валидации ШГР на основе большего количества ВВП [4] продемонстрировали более чем в десять раз меньшую долю объясненной дисперсии. Эти результаты согласуются с данными о том, что применение ШГР для этнических групп, отличных от той, для которой разработана ШГР, могут приводить к другим результатам. Есть данные, что экстраполяция результатов GWAS, полученных на одной этнической группе, при разработке ШГР для других этнических групп, может приводить к неверным оценкам [8]. Так, в исследовании Ehret GB, et al. (2016) продемонстрировано, что ассоциации ВВП с АГ различны для представителей выборок европейского и неевропейского (южноазиатского, восточноазиатского и африканского) происхождения [13]. Показано, что точность генетических предсказаний для разных популяций снижается практически линейно по мере увеличения различий между выборками [17].

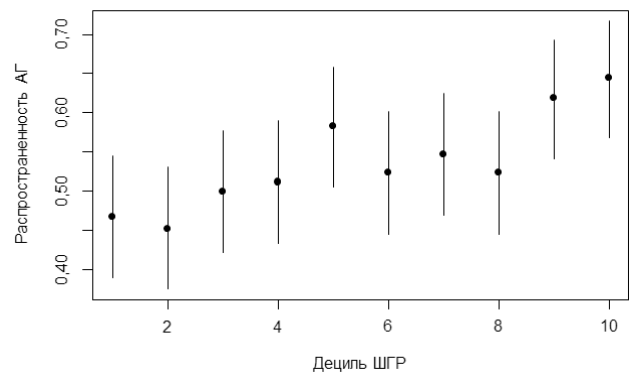


Рис. 4 Распространенность АГ для объединенной ШГР на основе 404 ВВП.

Примечание: АГ — артериальная гипертензия, ВВП — вариант нуклеотидной последовательности, ШГР — шкала генетического риска.

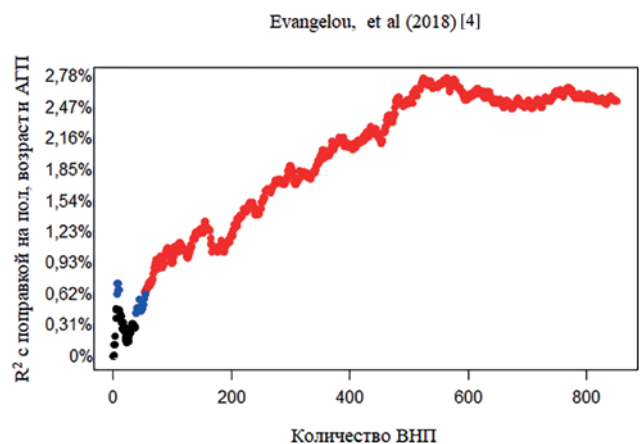


Рис. 5 Результаты линейной регрессии для ассоциации уровня САД со значениями ШГР (Evangelou E, et al, 2018) [4] после включения большего количества ВВП.

Примечание: для корреляции Пирсона между САД и значением ШГР: черный — $p > 0,05$, синий — $p < 0,05$, красный — $p < 0,005$. АГП — прием антигипертензивных препаратов, ВВП — вариант нуклеотидной последовательности, САД — систолическое артериальное давление, ШГР — шкала генетического риска, R^2 — коэффициент детерминации. Цветное изображение доступно в электронной версии журнала.

Для исследований, посвященных разработке ШГР, отсутствует стандартизированный набор оцениваемых на основании ШГР параметров, что может затруднять сравнение результатов разных исследований между собой. Только в одном из 4-х исследований (ШГР которых были выбраны для валидации в рамках нашего исследования) авторы оценивают распространенность АГ между нижним и верхним децилями [12]. При этом авторы получают значимые различия, однако интерпретация затруднительна в силу существенных различий размеров выборки [5]. Отсутствует единообразие в оценке метрик качества/предсказательной способности ШГР, с этой целью часто оценивают доли объясненной дисперсии. Ее увеличение при созда-

нии объединенной ШГР в рамках данного исследования согласуется с фактом, что предсказательная способность ШГР возрастает при увеличении количества ВВП, включенных в ШГР. Так, для одной и той же когорты было продемонстрировано, что увеличение количества ВВП с 29 [18] до 858 [19] позволяло увеличить предсказательную ценность ШГР. В настоящем исследовании R^2 увеличился с 0,9 до 2,5% при увеличении количества включенных ВВП с 28 до 852. Однако стоит отметить, что подобное конструирование ШГР имеет ограничения [5]: р-значения исходных ШГР зависели от размеров выборок, использованных для построения этих ШГР; коэффициенты исходных ШГР могли быть поправлены на структуру сцепленности ВВП.

На основании проведенной валидации ШГР с созданием собственной ШГР, наибольшую предсказательную способность имела ранее разработанная ШГР, включающая наибольшее количество ВВП. Это можно объяснить тем, что объединение необходимых статистических параметров (β -коэффициенты и p -value) из нескольких ШГР

с целью создания собственных ШГР не является оптимальным подходом. Более результативным является применение готовых ШГР (после их валидации), включающих наибольшее количество ВВП, либо создание ШГР на основе данных GWAS, в т.ч. на представителях российской популяции.

Заключение

ШГР АГ, разработанные на европейских выборках, не рекомендуется использовать у представителей российской популяции без предварительной валидации. Прогностическая ценность ШГР может быть улучшена за счет увеличения количества ВВП, включенных в ШГР. Для создания собственных ШГР не рекомендуется объединять статистические параметры (β -коэффициенты и p -value) из различных ШГР.

Отношения и деятельность. Научно-исследовательская работа "Разработка интегрированных систем прогнозирования в персонализированной медицине". Рег. № 121021700364-1.

Литература/References

- Williams B, Mancia G, Spiering W, et al. 2018 Practice Guidelines for the Management of Arterial Hypertension of the European Society of Hypertension and the European Society of Cardiology: ESH/ESC Task Force for the Management of Arterial Hypertension. *J Hypertens.* 2018; 36(12):2284-309. doi:10.1097/HJH.0000000000001961.
- Oparil S, Acelajado MC, Bakris GL, et al. Hypertension. *Nat Rev Dis Primers.* 2018;4:18014. doi:10.1038/nrdp.2018.14.
- Giri A, Hellwege JN, Keaton JM, et al. Trans-ethnic association study of blood pressure determinants in over 750,000 individuals. *Nat Genet.* 2019;51(1):51-62. doi:10.1038/s41588-018-0303-9.
- Evangelou E, Warren HR, Mosen-Ansorena D, et al. Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. *Nat Genet.* 2018;50(10):1412-25. doi:10.1038/s41588-018-0205-x.
- Choi SW, Mak TSH, O'Reilly PF. Tutorial: a guide to performing polygenic risk score analyses. *Nat Protoc.* 2020;15(9):2759-72. doi:10.1038/s41596-020-0353-1.
- Vaura F, Kauko A, Suvila K, et al. Polygenic risk scores predict hypertension onset and cardiovascular risk. *Hypertension.* 2021; 77(4):1119-27. doi:10.1161/HYPERTENSIONAHA.120.16471.
- Abu-El-Haija A, Reddi HV, Wand H, et al. The clinical application of polygenic risk scores: A points to consider statement of the American College of Medical Genetics and Genomics (ACMG). *Genet Med.* 2023;25(5):100803. doi:10.1016/j.gim.2023.100803.
- Martin AR, Gignoux CR, Walters RK, et al. Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. *Am J Hum Genet.* 2017;100(4):635-49. doi:10.1016/j.ajhg.2017.03.004.
- Morales J, Welter D, Bowler EH, et al. A standardized framework for representation of ancestry data in genomics studies, with application to the NHGRI-EBI GWAS Catalog. *Genome Biol.* 2018; 19(1):21. doi:10.1186/s13059-018-1396-2.
- Sakaue S, Hirata J, Kanai M, et al. Dimensionality reduction reveals fine-scale structure in the Japanese population with consequences for polygenic risk prediction. *Nat Commun.* 2020;11(1):1569. doi:10.1038/s41467-020-15194-z.
- Nauchno-organizatsionnyi komitet proekta ESSE-RF. Epidemiology of cardiovascular diseases in different regions of Russia (ESSE-RF). The rationale for and design of the study. *Profilakticheskaya Meditsina.* 2013;16(6):25-34. (In Russ.) Научно-организационный комитет проекта ЭССЕ-РФ. Эпидемиология сердечно-сосудистых заболеваний в различных регионах России (ЭССЕ-РФ). Обоснование и дизайн исследования. *Профилактическая медицина.* 2013;16(6):25-34.
- Ehret GB, Munroe PB, Rice KM, et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature.* 2011;478(7367):103-9. doi:10.1038/nature10405.
- Ehret GB, Ferreira T, Chasman DI, et al. The genetics of blood pressure regulation and its target organs from association studies in 342,415 individuals. *Nat Genet.* 2016;48(10):1171-84. doi:10.1038/ng.3667.
- Warren HR, Evangelou E, Cabrera CP, et al. Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk. *Nat Genet.* 2017;49(3):403-15. doi:10.1038/ng.3768.
- Ramensky VE, Ershova AI, Zaichenok M, et al. Targeted Sequencing of 242 Clinically Important Genes in the Russian Population From the Ivanovo Region. *Front Genet.* 2021;12. doi:10.3389/FGENE.2021.709419/FULL.
- van der Auwera G, O'Connor BD. Genomics in the Cloud: Using Docker, GATK, and WDL in Terra. O'Reilly Media, Inc. 2020. ISBN: 9781491975190.
- Scutari M, Mackay I, Balding D. Using Genetic Distance to Infer the Accuracy of Genomic Prediction. *PLoS Genet.* 2016;12(9): e1006288. doi:10.1371/journal.pgen.1006288.
- Fava C, Sjögren M, Montagnana M, et al. Prediction of blood pressure changes over time and incidence of hypertension by a genetic risk score in swedes. *Hypertension.* 2013;61(2):319-26. doi:10.1161/HYPERTENSIONAHA.112.202655.
- Giontella A, Sjögren M, Lotta LA, et al. Clinical Evaluation of the Polygenic Background of Blood Pressure in the Population-Based Setting. *Hypertension.* 2021;77:169-77. doi:10.1161/hypertensionaha.120.15449.