

Биобанкирование в стационаре многопрофильного научного медицинского центра как потенциал для широкого спектра научных исследований. Часть II. Особенности и первые результаты формирования аннотированной коллекции биоматериала

Копылова О.В.¹, Ершова А.И.¹, Покровская М.С.¹, Мешков А.Н.¹, Борисова А.А.¹, Ефимова И.А.¹, Серебрянская З.З.¹, Пустеленин А.В.¹, Юзьков Ю.В.², Гарбузова Е.В.¹, Величко А.М.¹, Сайфутдинов Ф.Р.¹, Драпкина О.М.¹

¹ФГБУ "Национальный медицинский исследовательский центр терапии и профилактической медицины" Минздрава России. Москва; ²ООО "СБС Контакт". Москва, Россия

В связи с развитием генетических и других видов исследований в сфере трансляционной медицины требуются коллекции, содержащие значительное количество биообразцов и большие массивы данных. Поскольку в процессе госпитализации чаще всего проводятся углубленное обследование пациентов, перспективным является создание коллекций, опирающихся на данные электронных историй болезни (ЭИБ). В связи с этим был инициирован проект "Сплошное биобанкирование в стационаре ФГБУ "НМИЦ ТПМ" Минздрава России", направленный на проведение сплошного биобанкирования крови и ее производных от пациентов стационарных отделений и создание единой системы, объединяющей информацию по биообразцам, клинические и другие виды данных, имеющихся в ЭИБ пациентов.

Цель. Описать коллекцию биообразцов, собранных в рамках научного проекта "Сплошное биобанкирование в стационаре ФГБУ "НМИЦ ТПМ" Минздрава России" на август 2023г, и проанализировать особенности создания обширной аннотации биообразцов пациентов в рамках данного проекта.

Материал и методы. Набор пациентов осуществляется на базе стационарных отделений ФГБУ "НМИЦ ТПМ" Минздрава России, только в случае подписания пациентом информированного согласия. После взятия крови биоматериал поступает в биобанк, где проводится пробоподготовка, размещение на хранение сыворотки, плазмы, цельной крови и занесение данных о каждом биообразце в специальную систему менеджмента биообразцов биобанка "FreezerPRO". Одновременно с этим собираются клинические данные о пациенте в рамках ЭИБ, которые объединены в медицинской информационной системе "Медиалог".

Результаты. Разработано специальное "Программное обеспечение для обмена данными между медицинской информационной системой и биобанком", которое объединяет информацию о пациентах, поступающую в медицинскую информационную систему "Медиалог" при заполнении ЭИБ, и систему менеджмента биообразцов биобанка "FreezerPro", содержащую подробную инфор-

мацию о биообразцах. На 25 августа 2023г собран биоматериал 3996 пациентов, который сопровождается подробной клинической аннотацией. В частности, доступны данные об основном и сопутствующих диагнозах (для 100% пациентов), данные лабораторных анализов (100%), инструментальных методов обследования, в частности, электрокардиограммы (100%), эхокардиографии (76,8%), суточного мониторирования электрокардиограммы по Холтеру (59,8%), эзофагогастродуоденоскопии (27,3%) и др. Наиболее частыми нозологиями явились гипертоническая болезнь (76,0%), атеросклероз (69,9%), ишемическая болезнь сердца (52,9%), нарушения ритма сердца (50,6%), сахарный диабет (19,2%). У 25,0% пациентов в анамнезе стентирование артерий.

Заключение. Сплошное биобанкирование в рамках стационара многопрофильного научного медицинского центра позволяет оперативно и качественно осуществлять набор биоматериала и ассоциированных данных. Собрана коллекция биообразцов с разносторонней и подробной клинической аннотацией. Проведен первичный анализ состава коллекции, показано, что она может быть использована для широкого спектра научных исследований.

Ключевые слова: биобанк, биобанкирование, ассоциированные данные, электронные истории болезни, клиническая аннотация.

Отношения и деятельность. Исследование выполнено в рамках государственного задания "Разработка интегрированных систем прогнозирования в персонализированной медицине" № 12102700364-1.

Поступила 18/10-2023

Рецензия получена 14/11-2023

Принята к публикации 22/11-2023



*Автор, ответственный за переписку (Corresponding author):

e-mail: sivoksana@yandex.ru

[Копылова О.В. — с.н.с., к.м.н., ORCID: 0000-0001-5397-5387, Ершова А.И. — д.м.н., зам. директора по фундаментальной науке, руководитель лаборатории клиномики, ORCID: 0000-0001-7989-0760, Покровская М.С. — к.б.н., руководитель лаборатории "Банк биологического материала", ORCID: 0000-0001-6985-7131, Мешков А.Н. — д.м.н., директор Института персонализированной терапии и профилактики, ORCID: 0000-0001-5989-6233, Борисова А.А. — ведущий инженер лаборатории "Банк биологического материала", ORCID: 0000-0003-4020-6647, Ефимова И.А. — ведущий эксперт лаборатории "Банк биологического материала", ORCID: 0000-0002-3081-8415, Серебрянская З.З. — лаборант лаборатории "Банк биологического материала", ORCID: 0000-0002-8691-1395, Пустеленин А.В. — аналитик группы по анализу качества медицинской помощи, ORCID: 0000-0003-3150-5194, Юзьков Ю.В. — администратор баз данных, ORCID: 0009-0001-6943-6588, Гарбузова Е.В. — лаборант-исследователь лаборатории клиномики, ORCID: 0009-0002-3184-7573, Величко А.М. — врач-ординатор, ORCID: 0009-0006-1505-605X, Сайфутдинов Ф.Р. — врач-ординатор, ORCID: 0000-0002-6257-6529, Драпкина О.М. — д.м.н., профессор, директор, ORCID: 0000-0002-4453-8430].

Для цитирования: Копылова О.В., Ершова А.И., Покровская М.С., Мешков А.Н., Борисова А.Л., Ефимова И.А., Серебрянская З.З., Пустеленин А.В., Юзков Ю.В., Гарбузова Е.В., Величко А.М., Сайфутдинов Ф.Р., Драпкина О.М. Биобанкирование в стационаре многопрофильного научного медицинского центра как потенциал для широкого спектра научных исследований. Часть II. Особенности и первые результаты формирования аннотированной коллекции биоматериала. *Кардиоваскулярная терапия и профилактика*. 2023;22(11):3799. doi:10.15829/1728-8800-2023-3799. EDN VFFYDN

Данные о предыдущей публикации: Борисова А.Л., Копылова О.В., Покровская М.С., Ефимова И.А., Пустеленин А.В., Ершова А.И., Драпкина О.М. Биобанкирование в стационаре многопрофильного научного медицинского центра как потенциал для широкого спектра научных исследований. Часть I. Организационно-методические аспекты. *Кардиоваскулярная терапия и профилактика*. 2023;22(11):3749. doi:10.15829/1728-8800-2023-3749. EDN BGDYIT

Biobanking in the hospital of a multidisciplinary research medical center as a potential for a wide research range. Part II. Specifics and first results of developing a described collection of biomaterial

Kopylova O. V.¹, Ershova A. I.¹, Pokrovskaya M. S.¹, Meshkov A. N.¹, Borisova A. L.¹, Efimova I. A.¹, Serebryanskaya Z. Z.¹, Pustelenin A. V.¹, Yuzkov Yu. V.², Garbuzova E. V.¹, Velichko A. M.¹, Sayfutdinov F. R.¹, Drapkina O. M.¹

¹National Medical Research Center for Therapy and Preventive Medicine. Moscow; ²ООО SBS Contact. Moscow, Russia

In connection with the development of genetic and other types of research in translational medicine, collections containing a significant number of biospecimens and large data sets are required. Since in-depth examination of patients is most often carried out during hospitalization, the creation of collections based on electronic medical records (EMR) data is promising. In this regard, the project "Total biobanking in the hospital of the National Medical Research Center for Therapy and Preventive Medicine" was initiated, aimed at total biobanking of blood and related products from patients in inpatient departments and the creation of a unified system that combines information on biosamples, clinical and other types of data.

Aim. To describe the collection of biospecimens collected within the project "Total biobanking in the hospital of the National Medical Research Center for Therapy and Preventive Medicine" for August 2023, and to analyze the process of biospecimen description.

Material and methods. Recruitment of patients is carried out on the basis of inpatient departments of the National Medical Research Center for Therapy and Preventive Medicine, only if the patient signs informed consent. After blood is taken, the biomaterial enters the biobank, where sample preparation and storage is carried out, as well as data about each biosample is entered into a special FreezerPRO system. At the same time, clinical data about the patient is collected using EMR, which are combined in the Medialog medical information system.

Results. A special software for data exchange between the medical information system and the biobank has been developed, which combines information about patients entering the Medialog medical information system and the FreezerPro biospecimen management system, which contains detailed information about biospecimens. As of August 25, 2023, biomaterial from 3996 patients was collected, which was accompanied by a detailed clinical summary. In particular, data is available on diagnoses (for 100% of patients), laboratory test data (100%), paraclinical data, in particular, electrocardiography (100%), echocardiography (76,8%), 24-hour Holter monitoring (59,8%), esophagogastroduodenoscopy (27,3%), etc. The most common diseases were hypertension (76,0%), atherosclerosis (69,9%), coronary artery disease (52,9%), arrhythmias (50,6%), diabetes (19,2%). In addition, 25,0% of patients had prior stenting.

Conclusion. Total biobanking within the hospital of a multidisciplinary research medical center allows for the rapid and high-quality collection of biomaterial and related data. A collection of biospecimens with comprehensive and detailed clinical annotations has been collected.

A primary analysis of the collection was carried, which showed that it can be used for a wide range of studies.

Keywords: biobank, biobanking, associated data, electronic medical records, clinical abstract.

Relationships and Activities. The work was carried out within the state acknowledgement "Development of integrated prediction systems in personalized medicine" № 121021700364-1.

Kopylova O.V.* ORCID: 0000-0001-5397-5387, Ershova A.I. ORCID: 0000-0001-7989-0760, Pokrovskaya M.S. ORCID: 0000-0001-6985-7131, Meshkov A.N. ORCID: 0000-0001-5989-6233, Borisova A.L. ORCID: 0000-0003-4020-6647, Efimova I.A. ORCID: 0000-0002-3081-8415, Serebryanskaya Z.Z. ORCID: 0000-0002-8691-1395, Pustelenin A.V. ORCID: 0000-0003-3150-5194, Yuzkov Yu. V. ORCID: 0009-0001-6943-6588, Garbuzova E.V. ORCID: 0009-0002-3184-7573, Velichko A.M. ORCID: 0009-0006-1505-605X, Sayfutdinov F.R. ORCID: 0000-0002-6257-6529, Drapkina O.M. ORCID: 0000-0002-4453-8430.

*Corresponding author:

sivoksana@yandex.ru

Received: 18/10-2023

Revision Received: 14/11-2023

Accepted: 22/11-2023

For citation: Kopylova O.V., Ershova A.I., Pokrovskaya M.S., Meshkov A.N., Borisova A.L., Efimova I.A., Serebryanskaya Z.Z., Pustelenin A.V., Yuzkov Yu. V., Garbuzova E.V., Velichko A.M., Sayfutdinov F.R., Drapkina O.M. Biobanking in the hospital of a multidisciplinary research medical center as a potential for a wide research range. Part II. Specifics and first results of developing a described collection of biomaterial. *Cardiovascular Therapy and Prevention*. 2023;22(11):3799. doi:10.15829/1728-8800-2023-3799. EDN VFFYDN

Data about a previous publication: Borisova A.L., Kopylova O.V., Pokrovskaya M.S., Efimova I.A., Pustelenin A.V., Ershova A.I., Drapkina O.M. Biobanking in the hospital of a multidisciplinary research medical center as a potential for a wide research range. Part I. Organizational and methodological aspects. *Cardiovascular Therapy and Prevention*. 2023;22(11):3749. doi:10.15829/1728-8800-2023-3749. EDN BGDYIT

БД — база данных, ЛИС — лабораторная информационная система, МИС — медицинская информационная система, МКБ — международная классификация болезней, СМБ — система менеджмента биообразцов, ФВ ЛЖ — фракция выброса левого желудочка, ЭИБ — электронные истории болезни, eMERGE Network — The Electronic Medical Records and Genomics Network (Сеть электронных медицинских карт и геномики), GWAS — Genome-wide association studies (полногеномный поиск ассоциаций), ID — уникальный идентификационный номер, PGRM — phenotype-genotype reference map (референсная "фенотипически-генотипическая карта").

Ключевые моменты**Что известно о предмете исследования?**

- Организация масштабного качественного сбора биообразцов и данных является нетривиальной задачей и требует вовлечения большого количества ресурсов — финансовых, кадровых, организационных.
- Объединение информации из электронных историй болезни и данных о биообразцах, хранящихся в биобанке, является перспективным для создания обширных подробно аннотированных коллекций.

Что добавляют результаты исследования?

- Подход сплошного биобанкирования в рамках стационара многопрофильного научного медицинского центра позволяет оперативно и качественно осуществлять набор биоматериала и ассоциированных данных.
- Благодаря объединению данных электронных историй болезни и биобанка собираемая коллекция может быть использована для широкого спектра научных исследований, в т.ч. с применением омиксных технологий.

Key messages**What is already known about the subject?**

- Management of large-scale, high-quality collection of biospecimens and data is a sophisticated task and requires large number of financial, personnel and organizational resources.
- Combining information from electronic medical records and biospecimen data stored in a biobank holds promise for creating rich and well described collections.

What might this study add?

- Total biobanking within the hospital of a multidisciplinary research medical center allows for the rapid and high-quality collection of biomaterial and related data.
- By combining data from electronic medical records and a biobank, the collection can be used for a wide range of studies, including using omics technologies.

Введение

В настоящее время для обеспечения качества научных исследований все шире применяются технологии биобанкирования [1, 2]. Это необходимо для стандартизации и повышения воспроизводимости результатов научных исследований, связанных с биообразцами. Ценность биообразцов обусловлена прежде всего связанными с каждым образцом данными. Например, материал, полученный от пациента с очень редким заболеванием, имеет большую значимость, предоставляя возможности для дальнейшего изучения данной патологии. В то же время, если все имеющиеся о биообразце данные, в т.ч. диагноз пациента, неизвестны (например, были утеряны или незафиксированы при взятии биоматериала), хранение даже самых редких образцов практически теряет смысл (чаще всего биообразцы без сопутствующей информации подлежат утилизации). Ценность биообразцов от пациентов даже с довольно частыми заболеваниями возрастает по мере количества имеющихся сопряженных данных. В последние годы даже предлагается ввести новый термин "Биодатбанк" ("Biodatabank"), поскольку он более точно отражает суть работы и устройства современных биобанков¹.

В связи с этим важнейшим вопросом, который необходимо решить при сборе любой коллекции био-

материала, является обеспечение полноты и качества сопутствующих данных. Выделяют несколько основных видов данных, ассоциированных с биообразцами:

— Информация о биообразцах — время и способ забора, условия транспортировки, особенности пробоподготовки, температура и продолжительность хранения и т.п.;

— Клинические, социально-демографические и другие виды данных о человеке, от которого даный биообразец получен — пол, возраст, раса, диагноз, различные показатели лабораторных, инструментальных и других исследований;

— Данные, получаемые впоследствии уже в процессе хранения биообразца — при проведении научных исследований, например, при использовании части аликвот могут стать известны дополнительные данные о биомаркерах, в части работ также продолжается динамическое наблюдение за пациентами — появляются новые, в т.ч. проспективные данные [3].

Перспективным направлением в плане сбора большого количества достоверной и хорошо организованной информации является объединение данных биобанка, в которых хранится информация о биообразцах, и данных электронных историй болезни (ЭИБ), которые обычно хранятся в рамках медицинской информационной системы (МИС) учреждения [4]. ЭИБ представляют собой уникальный ресурс, в который, зачастую на протяжении длительного времени, стекается вся информация о здоровье пациента. Кроме того, данные продол-

¹ <https://www.centogene.com/about-us/biodatabank> (дата обращения 12.09.2023).

жают пополняться даже после забора биообразца, что позволяет проводить не только ретроспективный анализ, но и проспективное наблюдение.

Все больше исследовательских групп и консорциумов по всему миру обращают внимание на то, что объединение данных биобанка с ЭИБ предоставляет уникальные возможности, в т.ч. для проведения широкомасштабных исследований в области генетики, изучения фенотипических ассоциаций, а также имплементации результатов генетического анализа в реальную клиническую практику [5, 6]. Так, примером может служить биобанк Лозанны, в который поступают биообразцы из Университетского госпиталя [6]. С 2013г в данной больнице всем стационарным и амбулаторным пациентам предлагается сдать кровь для помещения ее в биобанк, а также, в случае исследования тканей по медицинским показаниям, берется согласие на биобанкирование тканей, остающихся после проведения основного клинического исследования. При этом пациентам предлагается подписать информированное согласие на широкое и неограниченное использование данных образцов, а также медицинской информации из ЭИБ в разнообразных научных исследованиях, включая анализ генома [6].

Подход объединения данных ЭИБ и биообразцов биобанка уже доказал свою научную и клиническую ценность в ряде исследований. В одном из медицинских центров Германии проанализированы истории болезни пациентов на предмет наличия фенотипа, схожего с фенотипом муковисцидоза. В результате этой работы были отобраны 400 человек, генетический анализ их биообразцов, хранящихся в биобанке, показал, что у 262 из них имеется муковисцидоз, который не был диагностирован ранее [7].

Известны также многоцентровые проекты, реализующие подобный подход. В 2007г был инициирован проект eMERGE Network (The Electronic Medical Records and Genomics Network, Сеть электронных медицинских карт и геномики) [8]. Основным фокусом проекта является объединение генетических данных, получаемых благодаря анализу биообразцов биобанков, и данных ЭИБ нескольких крупных медицинских центров США. С момента основания прошло несколько этапов данного проекта, в результате которых сеть eMERGE внесла значимый вклад в исследования в области клинической генетики, фенотипирования клинически значимых заболеваний, клинического аннотирования, возврата результатов генетических исследований в медицинскую практику и оценку клинических исходов. Интеграция данных позволяет разрабатывать шкалы генетического риска, изучать фенотипические и генотипические ассоциации, проводить исследования в сфере фармакогеномики, изучать особенности имплементации генетических данных в ЭИБ и многое другое [8].

Организация масштабного качественного сбора биообразцов и связанных с ними данных является нетривиальной задачей и требует вовлечения большого количества ресурсов — финансовых, кадровых, организационных. В ФГБУ "Национальный медицинский исследовательский центр терапии и профилактической медицины" Министерства здравоохранения Российской Федерации (ФГБУ "НМИЦ ТПМ" Минздрава России) инициирован проект "Сплошное биобанкирование в стационаре ФГБУ "НМИЦ ТПМ" Минздрава России", реализация которого стала возможной благодаря наличию Биобанка и применения технологии биобанкирования. Подробности организации процесса сплошного биобанкирования описаны в первой части настоящей статьи [9]. После планирования и проведения организационных мероприятий 8 ноября 2022г состоялся старт проекта, сбор биоматериала в настоящий момент продолжается. Целью настоящей статьи является описание коллекции биообразцов, собранных в рамках научного проекта "Сплошное биобанкирование в стационаре ФГБУ "НМИЦ ТПМ" Минздрава России" на август 2023г, и анализ особенностей создания обширной аннотации биообразцов пациентов в рамках данного проекта.

Материал и методы

Проект одобрен на заседании Независимого этического комитета при ФГБУ "НМИЦ ТПМ" Минздрава России. Набор пациентов проводится сплошным методом на базе стационара ФГБУ "НМИЦ ТПМ" Минздрава России. Критериями включения в исследование явились госпитализация в одно из следующих отделений стационара: терапевтическое, кардиологическое, неврологическое, отделение неотложной кардиологии, отделение рентгенохирургических методов диагностики и лечения, отделение хирургического лечения сложных нарушений ритма сердца и электрокардиостимуляции, отделение реанимации и интенсивной терапии. Критерии невключения: 1) госпитализация в отделение сосудистой хирургии, 2) отказ пациента от подписания информированного согласия и участия в исследовании. Подписание информированного согласия на биобанкирование осуществляется в приемном отделении. Содержание информированного согласия одобрено Независимым этическим комитетом и включает информацию о возможности длительного хранения биоматериала в деперсонализированном виде и возможность его использования для будущих научных исследований, в т.ч. изучения генома. В случае, если пациент отказывается от участия в проекте, биобанкирование не осуществляется.

Каждому пациенту автоматически в МИС "Медиа-лог" присваивается уникальный идентификационный номер (ID) пациента. Далее сотрудник соответствующего отделения стационара формирует пакет документов (направление-талон на биобанкирование, информированное согласие, набор пробирок для взятия крови) на каждого пациента. Во время пребывания в стационаре (чаще всего на следующее утро) процедурная медсестра

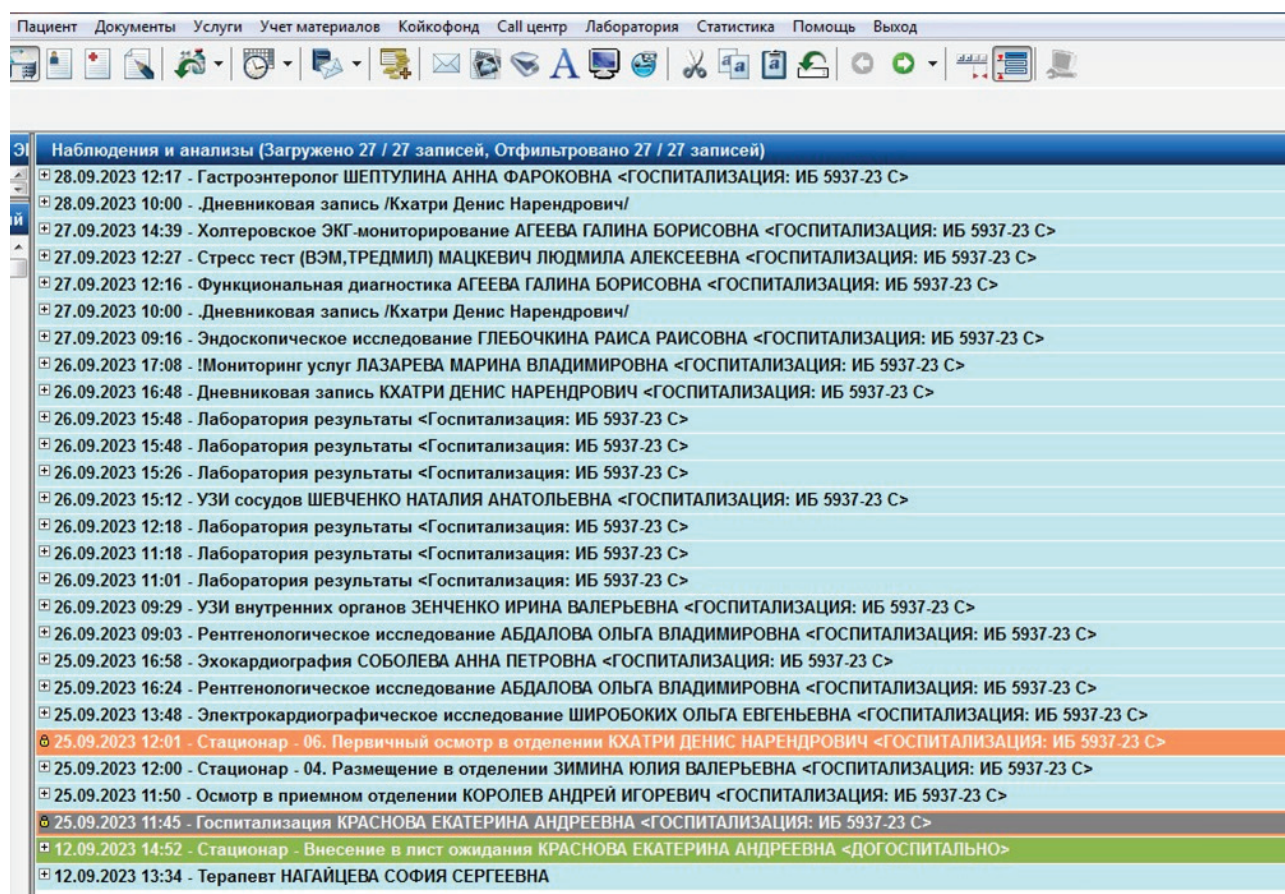


Рис. 1 Визуальный интерфейс МИС "Медиалог", содержащей информацию ЭИБ.

Примечание: МИС — медицинская информационная система, ЭИБ — электронные истории болезни.

осуществляет забор крови на биобанкирование (чаще всего совместно с другими запланированными анализами, отдельной венепункции не требуется). После этого сотрудник биобанка осуществляет транспортировку (в пределах учреждения), проводит пробоподготовку и размещение биообразцов на ответственное хранение в биобанк ФГБУ "НМИЦ ТПМ" Минздрава России, в т.ч. введение депersonализированной (с использованием ID пациента) информации в специализированную систему менеджмента биообразцов (СМБ) "FreezerPRO". На каждого пациента стандартно сохраняется 6 аликвот сыворотки (по 500 мкл), 3 аликвоты плазмы с калиевой солью этилендиаминтетраацетата (K_2 ЭДТА) (по 500 мкл) и 3 аликвоты плазмы с цитратом натрия (по 500 мкл), 1 образец цельной крови. Более подробно весь алгоритм биобанкирования отражен в первой части настоящей статьи [9].

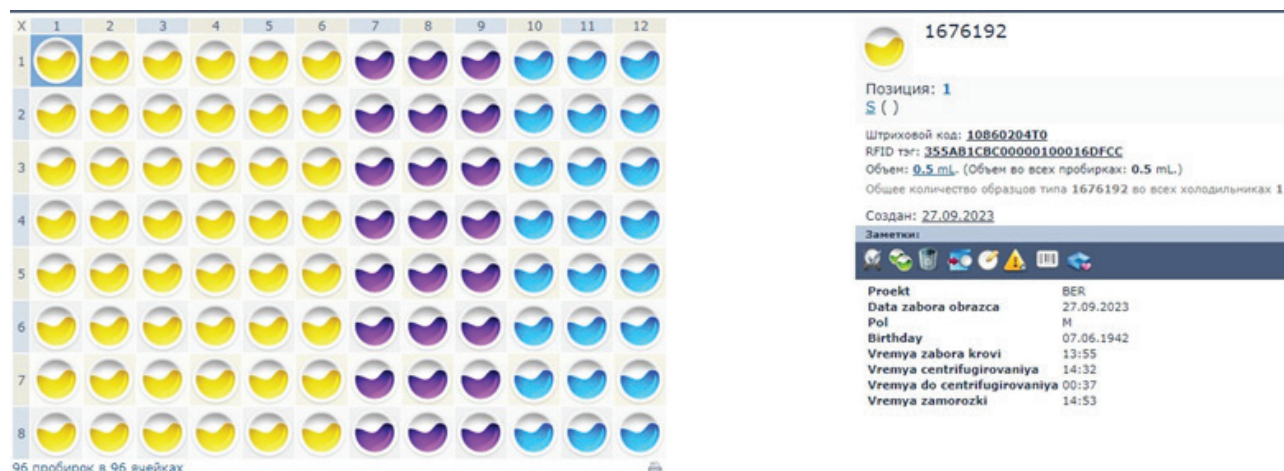
Сбор данных, составляющих аннотацию биообразцов, проводится при заполнении ЭИБ медицинским персоналом (например, данные физикального осмотра, анамнеза, эхокардиографии и т.п.). Лабораторные исследования, которые проводятся пациенту в рамках госпитализации, поступают в ЭИБ автоматически — непосредственно с анализаторов данные попадают в лабораторную информационную систему (ЛИС), а оттуда по специально разработанному программному интерфейсу автоматически передаются в МИС "Медиалог".

Анализ данных произведен на 25 августа 2023г, т.е. по прошествии 10 мес. от начала активной фазы проекта.

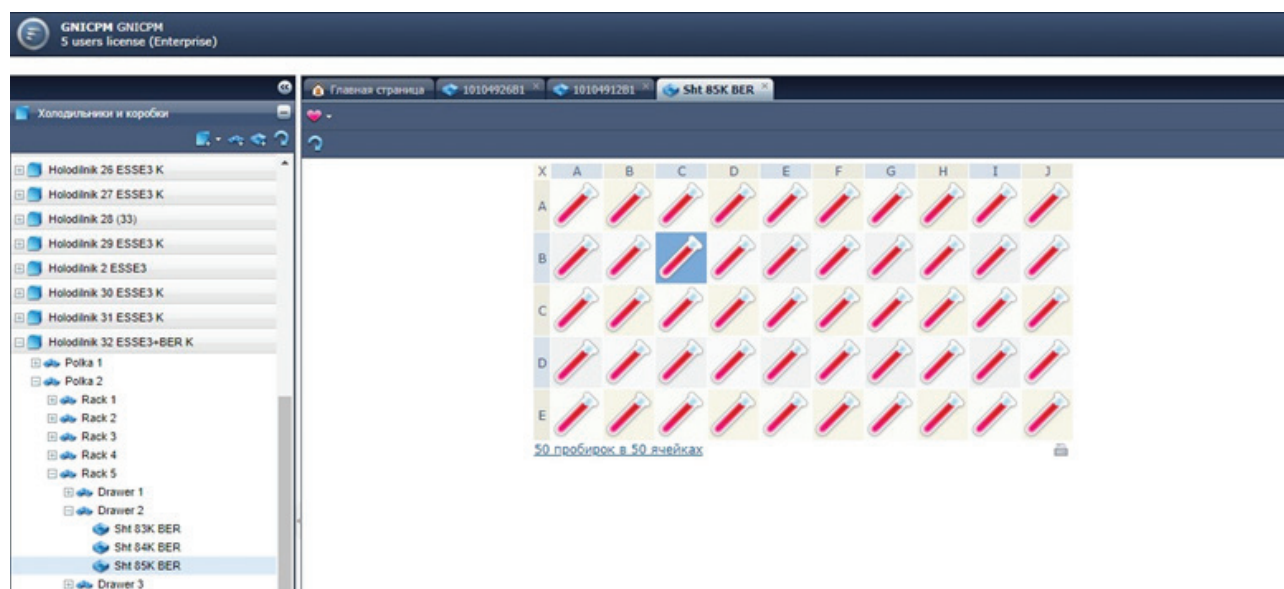
Статистическая обработка выполнена при помощи статистического пакета программ "Statistica 10.0". Данные представлены в виде абсолютных значений, процентных долей. Непрерывные показатели представлены в виде среднего значения и стандартного отклонения ($M \pm SD$).

Результаты и обсуждение

Благодаря сотрудничеству специалистов по информационным технологиям, биобанка, клиницистов и научных кадров было разработано специальное программное обеспечение (результат интеллектуальной деятельности "Программное обеспечение для обмена данными между медицинской информационной системой и биобанком", свидетельство о государственной регистрации программы для ЭВМ № 2023666507), которое объединяет данные о пациентах, формирующихся в МИС "Медиалог" при заполнении ЭИБ, и информацию о биообразцах, хранящуюся в СМБ "FreezerPro". Отметим, что созданная платформа обмена цифровыми данными необходима для работы исследователей-кураторов данного проекта для создания выборок пациентов по интересующим критериям и проведения научных работ с использованием биообразцов из биобанка.



А



Б

Рис. 2 Интерфейс СМБ "FreezerPRO" биобанка: А) описание каждого биообразца (желтый цвет (S) — сыворотка объемом 500 мкл, фиолетовый цвет — плазма с этилендиаминтетраацетатом объемом 500 мкл, голубой цвет — плазма с цитратом натрия объемом 500 мкл); Б) отображение точного местоположения каждого биообразца.

Примечание: СМБ — система менеджмента биообразцов. Цветное изображение доступно в электронной версии журнала.

Для обеспечения информационной безопасности и защиты данных платформа использует разграничение доступа к информации на уровнях доменной архитектуры, ролевой модели пользователей, а также средствами серверных систем управления базами данных (БД), хранящих информацию ЭИБ, лабораторных исследований и биообразцов. За счет использования децентрализованной модели хранения данных (МИС, ЛИС и БД биообразцов хранят информацию на разных серверах и системах управления БД) доступ к единому информационному массиву возможен только в соответствии с правами и ролями пользователей в системе и в объеме информации и данных, отведенных для роли. Предусмотрены следующие роли: врач стационарного отделения, сотрудник биобанка, сотрудник лаборатории, исследователь-куратор.

Все клинические, антропометрические, лабораторно-инструментальные, социально-демографические и другие виды данных о пациенте поступают в МИС "МедиаЛог". Визуальный интерфейс одной из страниц системы (рабочего стола врача в карте пациента) представлен на рисунке 1. Данные всех обследований, манипуляций, консультаций и других действий, касающихся пациента, поступают в организованной форме в систему. Все записи подтверждаются печатью врача (медицинского сотрудника), что крайне важно для обеспечения сохранности данных в неизменном виде.

Важным аспектом является сопряжение МИС "МедиаЛог" с ЛИС учреждения "Алиса". Благодаря этому все анализы, обрабатываемые лабораторией, в автоматическом режиме поступают в ЭИБ па-

Таблица 1

Наличие и средние значения
распространенных лабораторных показателей
в выборке пациентов проекта
"Сплошное биобанкирование в стационаре
ФГБУ "НМИЦ ТПМ" Минздрава России"

Показатель	Количество уникальных пациентов с доступными данными, n (%)	Среднее значение
Гемоглобин крови	3718 (93,0)	139±16 г/л
Глюкоза крови	3203 (80,2)	6,36±1,69 ммоль/л
Гликированный гемоглобин	861 (21,5)	6,4±1,3%
Креатинин	3566 (89,0)	84±24 мкмоль/л
Креатинин- фосфокиназа	3106 (77,7)	104±72 Ед/л
ХС ЛНП	3240 (81,1)	2,62±1,10 ммоль/л
Триглицериды	3262 (81,6)	1,40±0,82 ммоль/л
Липопротеин (а)	87 (2,2)	41,26±48,63 мг/дл

Примечание: ХС ЛНП — холестерин липопротеинов низкой плотности.

циента (отображаются в виде отдельных записей), хранящиеся в БД МИС с идентификатором пациента и идентификатором исследования.

Информация (например, демографические данные, основной диагноз согласно международной классификации болезней (МКБ), числовые показатели) МИС, ЛИС и биобанка, хранится в структурированном виде в реляционных БД в виде взаимосвязанных элементов и таблиц. При наличии соответствующего доступа необходимые данные могут быть выгружены в виде структурированных отчетов (таблиц), которые можно анализировать при помощи разных программ статистического анализа.

Часть информации (анамнез, полный клинический диагноз и т.п.) заполняется в специальных полях в текстовом формате. Одной из сложностей анализа данных в текстовом формате является разнообразие формулировок, которые используют разные специалисты. В частности, при составлении полного клинического диагноза использование формулировок МКБ зачастую затруднено в связи с тем, что данная классификация, прежде всего, служит целям медицинской статистики, и не всегда четко соотносится с реальной клинической практикой. В связи с этим МКБ обычно используется для кодирования основных диагнозов с целью дальнейшего анализа на уровне организации здравоохранения. При работе с пациентом в стационаре чаще всего используются свободные формулировки, которые позволяют более точно описать клинический статус пациента. К примеру, выражения "баллон-

ная ангиопластика со стентированием" — "БАП со стентированием", "транслюминальная баллонная ангиопластика со стентированием" — "ТЛБАП со стентированием", "транслюминальная баллонная коронарная ангиопластика со стентированием" — "ТБКА со стентированием" и др. означают одно и то же вмешательство, однако формулируются в диагнозах разными врачами по-разному. Кроме того, ручной ввод диагнозов может сопровождаться опечатками. Все это значительно затрудняет автоматизированный поиск по различным аспектам диагноза — в результаты поиска по заданным параметрам может попадать много "шумовых" данных, и, наоборот, часть пациентов упускается из виду. В связи с этим с целью унификации формулировки полных диагнозов клиническим и научным персоналом ФГБУ "НМИЦ ТПМ" Минздрава России проведена работа по созданию "Справочника диагнозов", содержащего диагнозы, наиболее часто используемые врачами стационара. Данные диагнозы сформулированы согласно новейшим актуальным клиническим рекомендациям — от врачей при добавлении диагноза в ЭИБ требуется использование именно включенных в справочник формулировок (МИС "Медиалог" позволяет добавлять готовые блоки формулировок одним нажатием клавиши компьютерной мыши, что исключает возможность опечаток, унифицирует формулировки, используемые врачами стационара, а также ускоряет процедуру ввода диагноза). При создании справочника учтены формулировки, имеющиеся в МКБ (10 и 11 пересмотр), Единой государственной информационной системе в сфере здравоохранения, актуальных версиях российских и международных клинических рекомендаций. Данный справочник интегрирован в МИС "Медиалог", периодически проводится его оптимизация с целью повышения удобства для пользователей, учета вновь возникающих клинических рекомендаций, касающихся рекомендуемых формулировок диагнозов. Использование "Справочника диагнозов" дает возможность повысить эффективность автоматизированного поиска пациентов с конкретной нозологией в МИС "Медиалог", снизить количество ошибок и опечаток (по сравнению с ручным вводом данных), унифицировать формулировки диагнозов среди врачей отделений разного профиля.

После взятия крови на биобанкирование проводится транспортировка, пробоподготовка и размещение биообразцов в биобанке, осуществляется занесение информации в СМБ "FreezerPRO", интерфейс которой представлен на рисунке 2. СМБ позволяет в наглядной форме видеть наличие каждой аликвоты и ее точное размещение (номер холодильника, номер полки, размещение на полке и др.).

Кроме выше указанной информации, БД СМБ "FreezerPRO" позволяет хранить дополнительную

информацию о биообразце — ID пациента в проекте, идентификационный номер записи в МИС "Медиалог", время забора биообразца, особенности пробоподготовки, температуру хранения и др.

Всего за данный период число пациентов, поступивших в стационар, составило 5385, количество пациентов, участвующих в биобанкировании в рамках данного научного проекта, составило 3996 (74,2%). Всего по данной выборке хранится 51948 образцов, из них сыворотки — 23976, плазмы с (K_2 ЭДТА) — 11988, плазмы с цитратом натрия — 11988, цельной крови — 3996.

Из включенных в проект пациентов 1989 (49,8%) — мужчины. Средний возраст пациентов составляет $63,9 \pm 13,3$ лет ($62,3 \pm 13,2$ у мужчин, $65,5 \pm 13,3$ у женщин). Среднее количество койко-дней, проведенных пациентами в стационаре, равнялось $5,7 \pm 3,1$.

МИС "Медиалог" дает возможность проанализировать статус курения включенных в проект пациентов: из госпитализированных в стационар курящими оказались 24,3%; в подгруппе мужчин курит 39,0%, в подгруппе женщин — 9,4%.

МИС "Медиалог" позволяет изучить диагнозы пациентов. Наиболее частыми патологиями явились: гипертоническая болезнь — 3037 (76,0%) человек (чел.), атеросклероз — 2794 (69,9%) чел., ишемическая болезнь сердца — 2115 (52,9%) чел., нарушения ритма сердца — 2022 (50,6%) чел., сахарный диабет — 766 (19,2%) чел., 1000 (25,0%) чел. когда-либо проводилось стентирование артерий.

Важным моментом является то, что стационарные пациенты чаще всего проходят достаточно углубленное обследование, результаты которого сохраняются в ЭИБ. В частности, значительному большинству пациентов проводятся различные лабораторные исследования: общий и биохимический анализы крови, коагулограмма, иммунологические исследования, общий анализ мочи и др. (таблица 1).

Данные об уровне гемоглобина и глюкозы крови натошак представлены для >90% чел.; о содержании в крови креатинина, холестерина липопротеинов низкой плотности — для >80% чел., уровня креатининфосфокиназы имеются в картах >75% чел., гликированного гемоглобина >20% чел. (таблица 1).

Объединение данных МИС "Медиалог" и СМБ "FreezerPRO" позволяет широко анализировать результаты инструментальных обследований. Всем пациентам при поступлении выполняется электрокардиография в 12 отведениях — 3996 (100%) чел. Значительному количеству пациентов в стационаре проводится эхокардиография, данные которой доступны для анализа. Так, показатель фракции выброса левого желудочка (ФВ ЛЖ) известен у 3068 (76,8%) чел., среднее значение ФВ ЛЖ составило $66,2 \pm 22,2\%$. Результаты суточного монитори-

рования электрокардиограммы по Холтеру имеются у 2389 (59,8%) чел. Данные рентгенографии органов грудной клетки доступны у 1209 (30,3%) чел., эзофагогастродуоденоскопии — у 1092 (27,3%) чел., мультиспиральной компьютерной томографии и магнитно-резонансной томографии различных областей — у 1525 (38,2%) чел.

ЭИБ предоставляют доступ к результатам консультаций различных специалистов. К примеру, заключения невролога имеются у 948 (23,7%) чел.; эндокринолога — у 392 (9,8%) чел.

Объединение данных ЭИБ, результатов лабораторных исследований и данных биобанка о биообразцах является крайне важным и позволяет исследователям, с одной стороны, проводить поиск необходимой клинической информации, связанной с биообразцом, а с другой — уточнять наличие биообразцов и их вид по каждому конкретному пациенту. Важным моментом является сохранение конфиденциальности данных пациентов, поскольку обмен информации происходит путем деперсонализации в виде использования ID пациента.

В мировой практике существует ряд инициатив, использующих подход объединения данных ЭИБ и биобанков. Одним из крупнейших проектов является Биобанк Великобритании (UK Biobank), который объединил биообразцы от >500 тыс. жителей Великобритании и сопряженную с ними социально-демографическую, фенотипическую, клиническую и другую информацию, в т.ч. содержащуюся в электронных медицинских картах Национальной системы здравоохранения. Проект обладает целым рядом уникальных особенностей: доступны генетические данные большинства лиц, чей биоматериал хранится в биобанке; наблюдение за когортой продолжается, в т.ч. благодаря сопряжению с данными в электронных медицинских картах, куда постоянно поступает новая клиническая информация о пациентах, что позволяет оценивать исходы в долгосрочной перспективе; проект открыт для ученых многих стран мира. Подобный подход позволяет изучать предикторы, закономерности развития, исходы широкого спектра заболеваний практически в любом медицинском направлении: кардиологии, онкологии, эндокринологии, пульмонологии и др. [10–13].

Изучение БД двух биобанков в Финляндии, сопряженных с ЭИБ, позволило разработать алгоритм для автоматического типирования хронической сердечной недостаточности на основании ФВ ЛЖ и лабораторных показателей. Алгоритм показал высокий уровень чувствительности и специфичности в отношении пациентов со сниженной и умеренно сниженной ФВ ЛЖ [14].

Объединение данных биобанков и ЭИБ предоставляет материал для проведения исследований, посвященных полногеномному поиску ассоциаций



Рис. 3 Два варианта реализации научных исследований от возникновения идеи до выдачи биообразцов и данных для проведения исследования.

(Genome-wide association studies, GWAS), в рамках которых изучают связи вариантов нуклеотидной последовательности, встречающихся в разных местах генома, с определенным состоянием или признаком, также известным как фенотип [15].

Перспективным является изучение фенотипически-генотипических ассоциаций (Phenome-wide association studies, PheWAS), в рамках которых проводится обратная процедура — поиск фенотипов, ассоциированных с конкретными вариантами нуклеотидной последовательности, в диапазоне тысяч фенотипов человека, так называемого "фенома" [15]. Bastarache L, et al. предлагают использовать референсную "фенотипически-генотипическую карту" (PGRM, phenotype-genotype reference map), которая составлена на основе 5879 генетических ассоциаций из 523 публикаций, посвященных GWAS. Фенотипы PGRM стандартизированы, что дает возможность обеспечить совместимость между биобанками [15].

Существуют разные подходы к сбору масштабных коллекций биообразцов. Один из подходов реализуется в рамках эпидемиологических исследований, которые подразумевают сбор биообразцов и данных от репрезентативных выборок населения (чаще всего подробная клиническая аннотация для каждого включенного участника отсутствует) с целью изучения распространенности заболеваний, их факторов риска и т.п. Подход с проведением сплошного биобанкирования также позволяет достаточно быстро собирать большое количество биообразцов, однако образцы поступают не от репрезентативной выборки, а от хорошо обследованных пациентов стационара, данные которых хранятся в организованных МИС, подобные выборки могут быть использованы с целью изучения ассоциаций

фенотип-генотип [8]. В рамках проекта "Сплошное биобанкирование в стационаре ФГБУ "НМИЦ ТПМ" Минздрава России" создана и продолжает пополняться ценная коллекция биообразцов стационарных пациентов, которая характеризуется качественными и подробными клиническими данными (в рамках текущей клинической практики).

Преимуществами использования данных ЭИБ являются точность получаемой информации (информация вводится медицинским персоналом, непосредственно участвующим в обследовании и лечении пациентов — отсутствует возможность возникновения ошибок при переносе данных), полнота данных (стационарные пациенты чаще всего хорошо обследованы), снижение времени от возникновения научной идеи до реализации исследования [4]. Так, при возникновении научной идеи существует два основных пути взаимодействия исследователей с биобанком (рисунок 3). Первый — это проспективный сбор и размещение на хранение биообразцов согласно теме исследования (подразумевает поиск пациентов, получение их согласия на участие в исследовании и постепенный набор коллекции биообразцов и данных). Второй — это подбор биообразцов из уже имеющихся в биобанке по заданным критериям, соответствующим задачам исследования (подразумевает изучение преаналитических и клинических данных). Безусловно, второй подход является значительно более быстрым и позволяет эффективно и в короткие сроки сформировать выборку пациентов и их биообразцов для проведения исследования.

В то же время при проведении сплошного биобанкирования имеется ряд ограничений: часть пациентов отказывается от участия в проекте (что не позволяет включить все 100% пациентов и может влиять на представительность выборки); инструментальные исследования проводятся разными операторами; длительность наблюдения и количество имеющихся данных может значительно различаться (одни пациенты проходили стационарное обследование и лечение лишь однократно, другие — несколько раз на протяжении последних лет); имеет место смещение выборки вследствие включения только стационарных пациентов.

Созданная коллекция может быть интересна при проведении различных исследований в сфере омиксных технологий, трансляционной и персонализированной медицины, направленных на создание шкал генетического риска, полногеномный поиск ассоциаций, поиск фенотипически-генотипических ассоциаций, поиск новых предиктивных и диагностических биомаркеров и др. В связи с большими перспективами использования подхода сплошного биобанкирования с последующим анализом биообразцов, хранящихся в биобанке, требуются дальнейшие исследования по оп-

тимизации алгоритмов анализа данных ЭИБ, в т.ч. изучения взаимосвязи геномов и феномов, а также возможности использования получаемой информации в повседневной клинической практике.

Заключение

Сплошное биобанкирование в рамках стационара ФГБУ "НМИЦ ТПМ" Минздрава России с объединением данных биобанка и информации из ЭИБ предоставляет уникальные возможности для создания ценной коллекции биообразцов, сопровождаемой качественными подробно описанными и структурированными данными. Менее чем

за год в проект включено ~4 тыс. пациентов. Таким образом, подход сплошного биобанкирования пациентов стационара позволяет оперативно и качественно осуществлять набор биоматериала, который, благодаря детализированной информации, хранящейся в ЭИБ, может быть использован для большого количества научных исследований различного профиля.

Отношения и деятельность. Исследование выполнено в рамках государственного задания "Разработка интегрированных систем прогнозирования в персонализированной медицине" № 12102700364-1.

Литература/References

1. Doludin YV, Borisova AL, Pokrovskaya MS, et al. Current best practices and biobanking recommendations. Russ Clin Lab Diagnostics. 2019;64(12):769-76. (In Russ.) Долудин Ю.В., Борисова А.Л., Покровская М.С. и др. Современные передовые практики и рекомендации по биобанкированию. Клиническая лабораторная диагностика. 2019;64(12):769-76. doi:10.18821/0869-2084-2019-64-12-769-776.
2. Pokrovskaya MS, Sivakova OV, Efimova IA, et al. Biobanking as a necessary tool for research in the field of personalized medicine in the scientific medical center. Per Med. 2019;16(6):501-9. doi:10.2217/pme-2019-0049.
3. Sivakova OV, Pokrovskaya MS, Metelskaya VA, et al. International rules for description of biospecimens are an important factor in improving the quality of researches. Profilakticheskaya Meditsina. 2019;22(6):95-99. (In Russ.) Сивакова О.В., Покровская М.С., Метельская В.А. и др. Международные правила описания биообразцов — важный фактор повышения качества научных исследований. Профилактическая медицина. 2019;22(6):95-99. doi:10.17116/profmed20192206295.
4. Cowie MR, Blomster JL, Curtis LH, et al. Electronic health records to facilitate clinical research. Clin Res Cardiol. 2017;106(1):1-9. doi:10.1007/s00392-016-1025-6.
5. Linder JE, Bastarache L, Hughey JJ, Peterson JF. The Role of Electronic Health Records in Advancing Genomic Medicine. Annu Rev Genomics Hum Genet. 2021;22:219-38. doi:10.1146/annurev-genom-121120-125204.
6. Barazzetti G, Bosisio F, Koutaissoff D, Spencer B. Broad consent in practice: lessons learned from a hospital-based biobank for prospective research on genomic and medical data. Eur J Hum Genet. 2020;28(7):915-24. doi:10.1038/s41431-020-0585-0.
7. Cassini T, Bastarache L, Zeng C, et al. A test of automated use of electronic health records to aid in diagnosis of genetic disease. Genet Med. 2023;25(12):100966. doi:10.1016/j.gim.2023.100966.
8. eMERGE Consortium. Lessons learned from the eMERGE Network: balancing genomics in discovery and practice. Hum Genet Genomics Adv. 2021;2(1):100018. doi:10.1016/J.XHGG.2020.100018.
9. Borisova AL, Kopylova OV, Pokrovskaya MS, et al. Biobanking in the hospital of a multidisciplinary research medical center as a potential for a wide research range. Part I. Organizational and methodological aspects. Cardiovascular Therapy and Prevention. 2023;22(11):3749. (In Russ.) Борисова А.Л., Копылова О.В., Покровская М.С. и др. Биобанкирование в стационаре многопрофильного научного медицинского центра как потенциал для широкого спектра научных исследований. Часть I. Организационно-методические аспекты. Кардиоваскулярная терапия и профилактика. 2023;22(11):3749. doi:10.15829/1728-8800-2023-3749.
10. Bešević J, Lacey B, Conroy M, et al. New Horizons: the value of UK Biobank to research on endocrine and metabolic disorders. J Clin Endocrinol Metab. 2022;107(9):2403-10. doi:10.1210/clinem/dgac407.
11. Conroy MC, Lacey B, Bešević J, et al. UK Biobank: a globally important resource for cancer research. Br J Cancer. 2023;128(4):519-27. doi:10.1038/s41416-022-02053-5.
12. Khera AV, Chaffin M, Wade KH, et al. Polygenic Prediction of Weight and Obesity Trajectories from Birth to Adulthood. Cell. 2019;177(3):587-96.e9. doi:10.1016/j.cell.2019.03.028.
13. Wang L, Fang R, Zhu M, et al. Integrated gene-based and pathway analyses using UK Biobank data identify novel genes for chronic respiratory diseases. Gene. 2021;767:145287. doi:10.1016/j.gene.2020.145287.
14. Vuori MA, Kiiskinen T, Pitkänen N, et al. Use of electronic health record data mining for heart failure subtyping. BMC Res Notes. 2023;16(1):208. doi:10.1186/s13104-023-06469-x.
15. Bastarache L, Denny JC, Roden DM. Phenome-Wide Association Studies. JAMA. 2022;327(1):75. doi:10.1001/jama.2021.20356.