

ОСОБЕННОСТИ СОЗДАНИЯ БАЗЫ ДАННЫХ БИОЛОГИЧЕСКИХ ОБРАЗЦОВ ДЛЯ БИОБАНКА

Максим Юрьевич Донников¹, Анна Владимировна Морозкина²,
Людмила Васильевна Коваленко³, Федор Федорович Иванов⁴,
Илья Михайлович Филиппов⁵, Артем Алексеевич Уколов⁶

^{1,2,3,4,5,6}Сургутский государственный университет, Сургут, Россия

¹donnikov@gmail.com, <https://orcid.org/0000-0003-0120-4163>

²morozkina_av@surgu.ru, <https://orcid.org/0009-0000-0547-4959>

³medsurdirector@gmail.com, <https://orcid.org/0000-0001-5708-7328>

⁴IvanovFF@surgu.ru, <https://orcid.org/0009-0004-5959-5923>

⁵partyneverkilled@mail.ru, <https://orcid.org/0009-0002-5577-0566>

⁶123456789z492@gmail.com, <https://orcid.org/0009-0002-9643-2029>

Аннотация. Цель – описание структуры и возможностей информационной модульной базы данных сопроводительной информации биологических образцов (крови и ее компонентов, геномной ДНК) для биобанка Медицинского института СурГУ.

Ключевые слова: база данных, биобанк, образцы биологического материала

Шифр специальности: 3.3.3. Патологическая физиология.

Для цитирования: Донников М. Ю., Морозкина А. В., Коваленко Л. В., Иванов Ф. Ф., Филиппов И. М., Уколов А. А. Особенности создания базы данных биологических образцов для биобанка // Вестник СурГУ. Медицина. 2023. Т. 16, № 2. С. 63–67. DOI 10.35266/2304-9448-2023-2-63-67.

Original article

FEATURES OF DEVELOPING A BIOBANK'S DATABASE OF BIOLOGICAL SAMPLES

Maksim Yu. Donnikov¹, Anna V. Morozkina², Lyudmila V. Kovalenko³,
Fedor F. Ivanov⁴, Ilya M. Filippov⁵, Artem A. Ukolov⁶

^{1,2,3,4,5,6}Surgut State University, Surgut, Russia

¹donnikov@gmail.com, <https://orcid.org/0000-0003-0120-4163>

²morozkina_av@surgu.ru, <https://orcid.org/0009-0000-0547-4959>

³medsurdirector@gmail.com, <https://orcid.org/0000-0001-5708-7328>

⁴IvanovFF@surgu.ru, <https://orcid.org/0009-0004-5959-5923>

⁵partyneverkilled@mail.ru, <https://orcid.org/0009-0002-5577-0566>

⁶123456789z492@gmail.com, <https://orcid.org/0009-0002-9643-2029>

Abstract. The study aims to describe a structure and functions of an information module database of biological samples associated information (blood and its components, genome DNA) for the biobank of the Medical Institute at Surgut State University.

Keywords: database, biobank, biological samples

Code: 3.3.3 Pathophysiology.

For citation: Donnikov M. Yu., Morozkina A. V., Kovalenko L. V., Ivanov F. F., Filippov I. M., Ukolov A. A. Features of developing a biobank's database of biological samples. *Vestnik SurGU. Meditsina*. 2023;16(2):63–67. DOI 10.35266/2304-9448-2023-2-63-67.

ВВЕДЕНИЕ

За последние годы с открытием биобанков в научно-исследовательских целях в Российской Федерации возникла необходимость в создании интегрированных баз сопроводительной информации для прове-

дения исследований с участием различных научных коллективов внутри государства и за его пределами.

База данных (БД) – совокупность данных, организованных по определенным правилам, включающая

общие принципы их описания, хранения и манипулирования [1]. В настоящее время нет общих стандартов для создания БД и систем управления данными биобанков, поэтому полнота и описание сильно варьируют в разных БД [2].

База данных может быть частью отдельного биобанка, однако в настоящее время появляются виртуальные БД, в которых хранится информация об образцах из реестра биобанков всей страны или даже нескольких государств [3]. В таком случае важно, чтобы базы создавались по одному принципу и могли быть конвертированы в общую БД.

Согласно классификации биологических БД по цели создания выделяют: научно-исследовательские, терапевтические и криминалистические БД. Научно-исследовательские подразделяются на популяционные, клинические и БД редких болезней [3].

В 2022 г. на базе Научно-образовательного центра (НОЦ) Медицинского института (МИ) Сургутского государственного университета (СурГУ) был создан научно-исследовательский нозологически-ориентированный (disease-oriented) биобанк для хранения образцов биоматериала (цельной крови и ее компонентов, геномной ДНК). В пяти коллекциях биобанка МИ СурГУ собраны образцы биоматериала доноров, перенесших коронавирусную инфекцию COVID-19, и доноров с постковидным синдромом, с патологией эндокринной системы, онкологическими заболеваниями (преимущественно рака желудка и колоректального рака), с патологией беременности и другими нарушениями репродукции, а также образцы биоматериала группы условно здоровых доноров (контрольная группа). Важной особенностью биологических коллекций является сбор и хранение образцов биоматериала представителей коренных народов ханты и манси, проживающих на территории Ханты-Мансийского автономного округа – Югры [4].

В связи с этим возникла необходимость в систематизации информации, ассоциированной с биологическими образцами, в виде создания специализированной БД.

Цель – описание разработки информационной модульной базы данных биологических образцов для биобанка Медицинского института СурГУ.

МАТЕРИАЛЫ И МЕТОДЫ

К марту 2023 г. включительно в коллекции биобанка НОЦ МИ СурГУ было собрано 4237 образцов крови и ее компонентов 534 доноров. До марта 2023 г. сопроводительная информация хранилась в таблицах Microsoft Excel, однако при «ручном» заполнении таблиц и отсутствии системы управления данными часто возникали опечатки и ошибки, а процесс внесения и корректировки данных был длительным и трудоемким.

Для создания автоматизированной системы (АС) управления данными было разработано программное обеспечение (ПО) в среде Microsoft Visual Studio-2019 на языке программирования C# – автоматизированная база данных (АБД) «Биобанк». При проектировании АС были использованы управляющие данные ГОСТ Р ИСО 20387-2021 «Биотехнология. Биобанкинг. Общие требования» и ряд положений Устава СурГУ, разработана инфологическая, функциональная и физическая модели с подробным описанием атрибутов. В качестве СУБД использовалась PostgreSQL. АБД «Биобанк» состоит из управляющего данными моду-

ля Main.cs и шести модулей, обеспечивающих функционирование. Интерфейс ПО соответствует принципу дружелюбности (user-friendly) для пользователя, как правило, специалиста с базовым образованием в области медицины и биологических наук, который обеспечен пошаговой инструкцией-руководством для работы с АБД «Биобанк» в нескольких режимах: добавления записей, удаления, редактирования, поиска. Предусмотрена возможность резервирования и восстановления БД. Для использования разработанного ПО достаточно операционной системы Windows 8/10.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

При анализе научной литературы подробное описание БД встречается намного реже, чем публикации по организации биобанков, подготовке и хранению образцов [2, 5, 6]. Чаще всего это связано с тем, что БД биобанка создается под определенный научно-исследовательский проект, а ее архитектура является объектом интеллектуального права и разглашению не подлежит.

При разработке структуры БД необходимо учитывать не только требования ученых конкретного исследования, но и возможность использования медико-биологической информации для будущих научных изысканий, например, в связи с появлением новых технологий анализа генов и геномов [7].

При проектировании БД биобанка НОЦ МИ СурГУ, по результатам анализа публикаций, посвященных данной проблеме [2, 4, 8, 9], были учтены основные принципы создания такого рода информационных ресурсов. Анализ примеров программного обеспечения с аналогичными задачами показал, что в отличие от коммерческих российских разработок, например, «Геном-эксперт. Биобанк» (ООО «Геном Эксперт»), и зарубежных – GeneLink (National Human Genome Research Institute – NHGRI, США), ISBER (International Society for Biological and Environmental Research), БД НОЦ МИ СурГУ нетребовательна к серверному оборудованию и, в отличие от БД НИИ медицинской генетики СО РАМН и БД CRUD-System Сибирского государственного медицинского университета, доступна для научных сотрудников СурГУ [2, 9].

ПО АБД «Биобанк» представляет собой приложение, позволяющее экспортировать данные в формате *.xls, осуществлять их редактирование/добавление/удаление и состоит из шести взаимосвязанных блоков (рис. 1).

Каждый из блоков выполняет свою функцию: AuthorizationForm.cs – авторизация пользователя; Form1.cs – главная форма; DBQueries.cs – обеспечение напрямую взаимодействующих с БД функций; DGVFuncs.cs – работающие с таблицами функции; DBDumpRestore – функции резервирования и восстановления данных; Etcetera.cs – функции вне категории.

Важными этапами работы с БД является этап описания образцов биоматериала и преаналитический этап для обеспечения условий качественного анализа результатов исследований [10]. Все сведения и биоматериалы, внесенные в биобанк на основании добровольного информированного согласия донора, подлежат анонимизации. БД могут быть и персонализированными [9, 11], но большинство авторов считают, что информация в научно-исследовательских БД должна быть анонимной, и случаи утечки данных не

Для оценки эффективности работы АБД «Биобанк» учет принятия образцов биоматериала на хранение ведется для каждой коллекции отдельно. За полтора года функционирования биорепоzitория НОЦ МИ СурГУ прирост количества образцов в коллекции «Патология беременности» происходил неравномерно, что связано с физическим перемещением первичной базы по сбору образцов в БУ ХМАО-Югры «Сургутский окружной клинический центр охраны материнства и детства», поэтому организационные проблемы сказались на снижении количества принимаемых образцов (рис. 3). Среднемесячный прирост образцов в 2022 г. составил 47,5 ед., в первом квартале 2023 г. снизился до 16, однако уже в марте увеличился до 36 образцов.

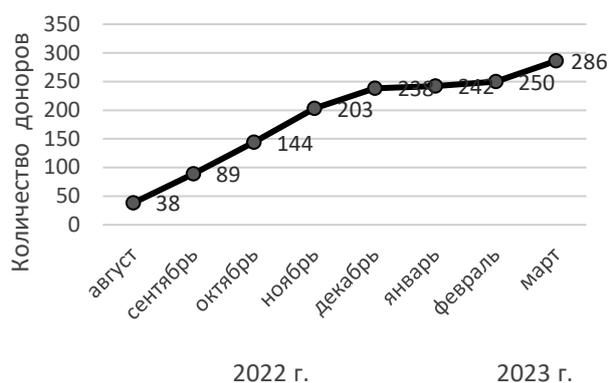


Рис. 3. Динамика поступления образцов в коллекцию «Патология беременности» АБД «Биобанк»

Поступление образцов биоматериала доноров с патологией сердечно-сосудистой системы в БД коллекции «КардиоГен» в 2023 г. было равномерным и составило 17 образцов ежемесячно (рис. 4).

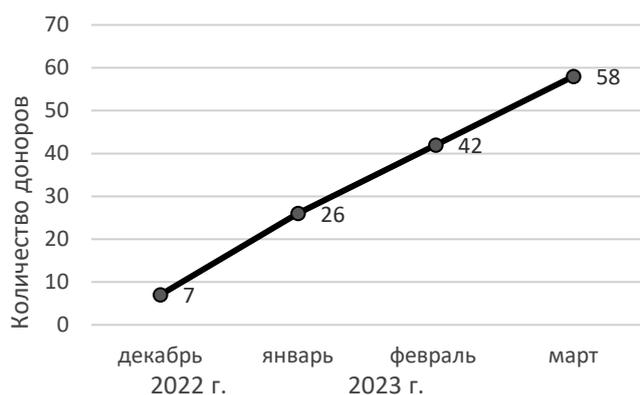


Рис. 4. Динамика поступления образцов в коллекцию «КардиоГен» АБД «Биобанк»

Эффективный сбор материалов и пополнение БД биобанка информацией об образцах биоматериала пациентов с онкологическими заболеваниями (коллекция «ОнкоПатология») определяется рядом организационных моментов, связанных с нестабильным состоянием потенциальных доноров. Увеличение поступивших образцов биоматериала в БД коллекции «ОнкоПатология» в начале 2023 г. по сравнению с 2022 г. обусловлено вовлечением врачей-онкологов в научно-исследовательскую работу и активным рекрутингом пациентов для пополнения биобанка биоматериалами (рис. 5).

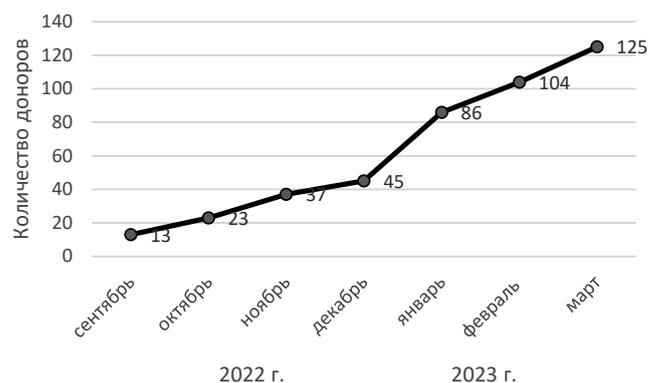


Рис. 5. Динамика поступления образцов в коллекцию «ОнкоПатология» АБД «Биобанк»

Так, если среднемесячный прирост образцов коллекции «ОнкоПатология» в 2022 г. составил 11 ед., то уже в первом квартале 2023 г. увеличился до 26 ед. ежемесячно, т.е. более чем в два раза.

ЗАКЛЮЧЕНИЕ

Таким образом, разработанная автоматизированная база данных для организации работы биобанка обеспечивает оперативное наполнение, надежное хранение и эффективное использование сопроводительной информации. Однако используемая актуальная версия АБД «Биобанк» пока не поддерживает сетевой вариант и не обеспечивает работу с темпоральными (временными) данными, что является предметом работы по улучшению текущей версии программного обеспечения.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Conflict of interest. The authors declare no conflict of interest.

СПИСОК ИСТОЧНИКОВ

1. Яценко Е. А., Иванов Ф. Ф., Егоров А. А. Курс лекций по базам данных. Сургут: Издат. центр СурГУ, 2014. Ч. 1. 96 с.
2. Буйкин С. В., Брагина Е. Ю., Конева Л. А. и др. Базы данных коллекций биологического материала: организация сопроводительной информации // Бюллетень сибирской медицины. 2012. Т. 11, № 1. С. 111–121.

REFERENCES

1. Yatsenko E. A., Ivanov F. F., Egorov A. A. Kurs lektii po bazam dannykh. Surgut: Surgut State University Publishing House; 2014. Pt. 1. 96 p. (In Russian).
2. Buikin S. V., Bragina E. Yu., Koneva L. A. et al. Databases of biological collection: Organization of associated information. *Bulletin of Siberian Medicine*. 2012;11(1):111–121. (In Russian).

3. Болтанова Е. С., Имекова М. П. Виды биологических баз данных (биобанков) // Вестник Томского государственного университета. Право. 2021. № 41. С. 136–148.
4. Калинин Р. С., Голева О. В., Илларионов Р. А. и др. Формирование биобанка в структуре научных и лечебно-диагностических учреждений и перспективы межрегиональной интеграции // Кардиоваскулярная терапия и профилактика. 2022. Т. 21, № 11. С. 3401.
5. Покровская М. С., Борисова А. Л., Метельская В. А. и др. Роль биобанкирования в организации крупномасштабных эпидемиологических исследований // Кардиоваскулярная терапия и профилактика. 2021. Т. 20, № 5. С. 2958.
6. ISO 20387:2018 Biotechnology – Biobanking – General requirements for biobanking. URL: <https://www.iso.org/standard/67888.html> (дата обращения: 20.03.2023).
7. Копылова О. В., Ершова А. И., Покровская М. С. и др. Популяционно-нозологический исследовательский биобанк «НМИЦ ТПМ»: анализ коллекций биообразцов, принципы сбора и хранения информации // Кардиоваскулярная терапия и профилактика. 2021. Т. 20, № 8. С. 3119.
8. Буйкин С. В., Брагина Е. Ю., Конева Л. А. Разработка структуры базы данных для биобанков // Якутский медицинский журнал. 2011. № 1. С. 71–73.
9. Гутор С. С., Энглевский Н. А., Прокудина Д. В. и др. Банк биологического материала: информационное сопровождение и физическое воплощение // Врач и информационные технологии. 2013. № 4. С. 31–39.
10. Сивакова О. В., Покровская М. С., Метельская В. А. и др. Международные правила описания биообразцов – важный фактор повышения качества научных исследований // Профилактическая медицина. 2019. Т. 22, № 6–2. С. 95–99.
11. Hallmans G., Vaught J. B. Best practices for establishing a biobank. In: Dillner, J., editor. *Methods in biobanking. Methods in molecular biology*. Vol 675. Totowa, NJ: Humana Press; 2011. p. 241–260.
12. Малеина М. Н. Правовой статус биобанка (банка биологических материалов человека) // Право. Журнал Высшей школы экономики. 2020. № 1. С. 98–117.
3. Boltanova E. S., Imekova M. P. The types of biological databases (biobanks). *Tomsk State University Journal of Law*. 2021;(41):136–148. (In Russian).
4. Kalinin R. S., Goleva O. V., Illarionov R. A. et al. Development of a biobank in the structure of scientific and diagnostic and treatment institutions and prospects for interregional integration. *Cardiovascular Therapy and Prevention*. 2022;21(11):3401. (In Russian).
5. Pokrovskaya M. S., Borisova A. L., Metelskaya V. A. et al. Role of biobanking in managing large-scale epidemiological studies. *Cardiovascular Therapy and Prevention*. 2021;20(5):2958. (In Russian).
6. ISO 20387:2018 Biotechnology – Biobanking – General requirements for biobanking. URL: <https://www.iso.org/standard/67888.html> (accessed: 20.03.2023).
7. Kopylova O. V., Ershova A. I., Pokrovskaya M. S. et al. Population-nosological research biobank of the National Medical Research Center for Therapy and Preventive Medicine: Analysis of biosamples, principles of collecting and storing information. *Cardiovascular Therapy and Prevention*. 2021;20(8):3119. (In Russian).
8. Buikin S. V., Bragina E. Yu., Koneva L. A. Working out of database structure for biobanks. *Yakut Medical Journal*. 2011;(1):71–73. (In Russian).
9. Gutor S. S., Englevsky N. A., Prokudina D. V. et al. Bank of biosamples: Information management and physical implementation. *Medical Doctor and IT*. 2013;(4):31–39. (In Russian).
10. Sivakova O. V., Pokrovskaya M. S., Metelskaya V. A. et al. International rules for description of biospecimens are an important factor in improving the quality of researches. *The Russian Journal of Preventive Medicine*. 2019;22(6–2):95–99. (In Russian).
11. Hallmans G., Vaught J. B. Best practices for establishing a biobank. In: Dillner, J., editor. *Methods in biobanking. Methods in molecular biology*. Vol 675. Totowa, NJ: Humana Press; 2011. p. 241–260.
12. Maleina M. N. Legal status of the biobank (bank of biological human material). *Law. Journal of the Higher School of Economics*. 2020;(1):98–117. (In Russian).

ИНФОРМАЦИЯ ОБ АВТОРАХ

М. Ю. Донников – кандидат медицинских наук, научный сотрудник.

А. В. Морозкина – кандидат биологических наук, научный сотрудник.

Л. В. Коваленко – доктор медицинских наук, профессор.

Ф. Ф. Иванов – кандидат технических наук, профессор.

И. М. Филиппов – студент.

А. А. Уколов – студент.

INFORMATION ABOUT THE AUTHORS

M. Yu. Donnikov – Candidate of Sciences (Medicine), Researcher.

A. V. Morozkina – Candidate of Sciences (Biology), Researcher.

L. V. Kovalenko – Doctor of Sciences (Medicine), Professor.

F. F. Ivanov – Candidate of Sciences (Engineering), Professor.

I. M. Filippov – Student.

A. A. Ukolov – Student.