



Метаданные в медицине

Валеева А. Х.-М.¹, Гомон Ю. М.^{1,2}, Белявская Н. М.³

¹ ФГБОУ ВО «Первый Санкт-Петербургский государственный медицинский университет имени академика И. П. Павлова», Санкт-Петербург, Российская Федерация

² СПбГБУЗ «Больница Святого Великомученика Георгия», Санкт-Петербург, Российская Федерация

³ ГБУЗ ЛО «Всеволожская клиническая межрайонная больница» врачебная амбулатория д. Кальтино, Ленинградская область, Российская Федерация

Аннотация

Наблюдательные исследования, имеющие решающее значение для понимания сложных явлений реального мира, часто страдают от плохой обнаруживаемости и воспроизводимости данных. Проект MINERVA (Метаданные для обнаружения данных и воспроизводимости исследований в наблюдательных исследованиях) в Европе был направлен на решение этих проблем путём разработки и внедрения надёжных стандартов и практик метаданных.

Ключевые слова: метаданные в медицине; медицинские информационные системы; MINERVA; наблюдательные исследования

Для цитирования: Валеева А. Х.-М., Гомон Ю. М., Белявская Н. М. Метаданные в медицине. *Реальная клиническая практика: данные и доказательства*. 2025;5(1):59-64. <https://doi.org/10.37489/2782-3784-myrd-069>. EDN: CVSRRD.

Поступила: 19.01.2025. **В доработанном виде:** 27.02.2025. **Принята к печати:** 14.03.2025. **Опубликована:** 31.03.2025.

Metadata in medicine

Altynai Kh-M. Valeeva¹, Yulia M. Gomon^{1,2}, Nadezhda M. Belyavskaya³

¹ First St. Petersburg State Medical University named after I. P. Pavlov, St. Petersburg, Russian Federation

² Hospital of the Holy Great Martyr George, Saint Petersburg, Russian Federation

³ Vsevolozhsk Clinical Interdistrict Hospital, medical outpatient clinic, village of Kaltino, Leningrad Region, Russian Federation

Abstract

Observational study, crucial for understanding complex real-world phenomena, often suffers from poor data detectability and reproducibility. The MINERVA (Metadata for Data Discovery and Research Reproducibility in Observational Studies) project in Europe aimed to address these issues by developing and implementing reliable metadata standards and practices.

Keywords: metadata in medicine; medical information systems; MINERVA; observational study

For citation: Valeeva AH-M, Gomon YuM, Belyavskaya NM. Metadata in medicine. *Real-World Data & Evidence*. 2025;5(1):59-64. <https://doi.org/10.37489/2782-3784-myrd-069>. EDN: CVSRRD.

Received: 19.01.2025. **Revision received:** 27.02.2025. **Accepted:** 14.03.2025. **Published:** 31.03.2025.

Метаданные в организации и классификации информации / Metadata in the organization and classification of information

Метаданные — структурированные данные, характеризующие информационный ресурс для целей его идентификации, поиска и управления им [1]. Метаданные помогают организовывать и классифицировать большие объёмы медицинских данных.

Метаданные можно разделить на следующие категории [3]:

- **административные:** регистр населения, регистр смертей, регистрация в системе здравоохранения, возмещение расходов на диагностические тесты или процедуры, административные требования в отношении медицинского обслуживания;
- **первичная медицинская помощь:** медицинские карты первичной медицинской помощи, записи о выдаче лекарств;
- **вторичная медицинская помощь:** записи о выписке из больницы, записи о стационарных

- и амбулаторных пациентах больницы, записи о выписке из неотложной помощи, записи о специализированной амбулаторной помощи;
- **реестры:** рождаемость, смертность, искусственное прерывание беременности, врождённые аномалии, онкологические заболевания, прививки, лекарственные препараты;
- **другое:** биобанк, сообщения о побочных реакциях на лекарственные препараты.

Метаданные мы можем получать из медицинской информационной системы (МИС). Они способствуют повышению эффективности работы медицинских учреждений, сокращению времени на административные задачи и улучшению качества медицинского обслуживания.

Однако на данный момент существует минимум две проблемы:

1. Много локальных МИС, в которых содержится информация о пациенте и не передаётся в другие МИС, из-за чего происходит проблема в том, что заболевания дублируются, неизвестно точ-

ное число заболевших. Это пагубно влияет на сбор статистических данных для метаанализа.

2. Не во всех больницах есть МИС. До сих пор в 2025 году в России существуют больницы с бумажным ведением документации, из-за чего информация о пациентах легко теряется и не учитывается при систематическом заборе метаданных.

Согласно традиционной модели оказания медицинских услуг, 39% рабочего времени врача уходит на оформление медицинских документов, а 50% — на поиск необходимой информации. В этой связи внедрение информационных систем (ИС) становится важным инструментом для оптимизации и экономии ресурсов. Особенно актуально это для России, где большой размер территории и недостаточная мощность сельских лечебных учреждений делают использование ИС ещё более востребованным [4].

Есть различные стратегии для централизации метаданных локальных систем [5].

Таблица 1. Стратегии для централизации метаданных локальных систем (по Gini R. с изменениями) [5]
Table 1. Strategies for centralizing metadata of local systems (according to Gini R. with changes) [5]

Стратегия	Описание	Преимущества	Недостатки
Локальных анализов	Каждое учреждение извлекает и анализирует данные локально, используя собственные программы	Такая стратегия полностью затрагивает все локальные метаданные, которые в централизованной системе могут не быть затронуты из-за малого количества случаев	Собственные программы не находятся под наблюдением централизованной системы, что в свою очередь может привести к доступу к недостоверным данным и неточной статистике. Это может привести к вводящим в заблуждение выводам и ложным прогнозам
Обмена сырыми данными	Данные извлекаются локально и передаются без анализа центральному партнёру для совместного анализа	Контролирование поступления данных, возможность локального анализа и поиска ошибок в конкретном учреждении. Такой подход даёт перспективу для дальнейших исследований на основе сырых данных. Также не стоит забывать, что совместный анализ даёт большой спектр мнений и возможность более объективного подхода	При транспортировке данных есть вероятность их потери, что может привести к искажению статистики. Необходимо дополнительное время для форматирования данных. При транспортировке на физическом носителе — повышение количества расходов
Использование модели общих данных с конкретными данными	Данные извлекаются, загружаются в общую модель данных и обрабатываются локально с единым программным обеспечением	Быстрая подготовка данных, нет необходимости в дополнительном форматировании. Упрощение хранения	Отсутствие гибкости информационной системы, что приводит к отсутствию индивидуального подхода к конкретному медицинскому учреждению. Долгий путь от медицинского учреждения до службы поддержки
Использование общей модели данных	Все локальные данные загружаются в общую модель данных независимо от протокола исследования	Сокращение времени загрузки данных, не нужно делать общий стандарт загрузки файлов и обучать персонал при переходе на новую систему	Некорректность подачи данных (нет общего формата данных), большая затрата времени на форматирование данных при анализе в центральной системе. Невозможность планирования хранения данных в силу разного формата файлов

Каждая стратегия требует разного времени для согласования протоколов и обработки данных.

Метаданные в сфере медицинских исследований / Metadata in medical research

Достоверные метаданные дают доступ к более качественным медицинским исследованиям.

В январе 2021 года в Европе был проведён тщательный поиск на вебсайтах ключевых организаций и консорциумов, обладающих опытом каталогизации баз данных здравоохранения и их метаданных. Из публикаций и каталожных инструментов, выявленных в ходе исследования, были извлечены рекомендации по сбору метаданных и соответствующие поля метаданных. На основе этой информации был составлен первоначальный список уникальных полей метаданных.

Затем были проведены структурированные интервью с представителями восьми организаций (FDA Sentinel Initiative, CNODES, IMI-EHDEN, IMI-ConcePTION, AsPEN, FAIRplus, Maelstrom, Action) имеющих опыт проведения фармакоэпидемиологических исследований с использованием нескольких источников данных или имеющих опыт работы с метаданными в сфере здравоохранения.

Интервью помогли уточнить список метаданных и определить дополнительные важные переменные метаданных, которые необходимо собрать. Они также помогли собрать дополнительные ресурсы и найти существующие инструменты для доступа к метаданным и их визуализации. Эксперты подчеркнули важность проведения тщательной проверки качества метаданных перед их вводом в инструменты визуализации, чтобы обеспечить их надлежащее функционирование и соответствие определениям.

Отзывы о первоначальном списке метаданных были получены в ходе опроса заинтересованных сторон и технического семинара в апреле 2021 года. На основе этих отзывов первоначальный список метаданных был пересмотрен, а ключевые метаданные для сбора были определены как приоритетные [6].

Метаданные и качество медицинского обслуживания / Metadata and quality of health care

Метаданные могут использоваться для оценки качества и целостности данных, что критически важно в медицинских исследованиях. Доступ к стандартному электронному набору полной и точной информации о метаданных может способствовать нахождению источников данных, подходящих для конкретного исследования, облегчать описание источников данных, которые планируется использовать в протоколе исследований

и способствовать оценке доказательной ценности результатов. Совместная целевая группа по большим данным руководителей агентств по лекарственным средствам и Европейского агентства по лекарственным средствам (HMA — EMA) рекомендовала «способствовать обнаружению данных посредством идентификации метаданных» в рамках своих рекомендаций.

Проект MINERVA (метаданные для обнаружения данных и воспроизводимости исследований в наблюдательных исследованиях) в Европе был направлен на решение проблем, связанных с отсутствием согласованных метаданных, препятствовавших возможности повторного использования информации во вторичных анализах, ограничивая общее влияние исходного исследования.

Проект был направлен на разработку и внедрение надёжных стандартов и практик метаданных.

На основе результатов исследования MINERVA и консультаций с сообществом ENCePP и другими заинтересованными сторонами, Европейское агентство по лекарственным средствам (EMA) разрабатывает электронный каталог, который будет содержать метаданные для облегчения поиска источников данных, т.е. для первоначального определения подходящих источников данных для изучения конкретного исследовательского вопроса, а также для оценки протоколов исследований и результатов исследований путём предоставления быстрого доступа к информации о пригодности источников данных [3].

ЕМА отвечает за научную оценку, контроль и мониторинг безопасности лекарств в Европейском Союзе (ЕС) для обеспечения их благоприятного соотношения риска и пользы. Система мониторинга безопасности лекарств начала формироваться с механизмов спонтанного сообщения о подозреваемых нежелательных реакциях и с тех пор расширилась, включая обширные наблюдательные исследования. С 2006 года, когда была введена система управления рисками, ЕМА инициировала создание Европейской сети центров по фармакоэпидемиологии и фармаконадзору (ENCEPP), которая была запущена в 2007 году.

Существующие годы в Европе сталкивались с проблемами из-за разнородности систем здравоохранения и ограниченных выборок данных. ENCePP была создана для повышения исследовательских мощностей в области фармакоэпидемиологии, разработки общих методологических стандартов и предложений по управлению исследованиями. ENCePP собрала 168 центров из 18 стран, 75% которых являются некоммерческими организациями.

Ежегодно обновляемые методологические рекомендации ENCePP стали справочным ресурсом в области фармакоэпидемиологии.

ENCePP успешно поддерживает сотрудничество в разнообразной среде, благодаря стремлению к совместным принципам прозрачности и независимости, а также активному участию участников в совместных мероприятиях и инициативах. За 10 лет ENCePP значительно способствовала оценке соотношения рисков и пользы лекарств в Европе, повысив прозрачность в области наблюдательных исследований и предоставив методологические рекомендации [7].

Метаданные и эффективный обмен информацией / Metadata and efficient information exchange

С помощью метаданных можно делиться информацией, которая полезна для обнаружения схожих проблем и быстрого их решения.

Важно, чтобы информация о метаданных постоянно обновлялась, чтобы работать с актуальными

данными. В идеале обновление информации надо проводить ежегодно или чаще для конкретных источников данных, если это будет сочтено необходимым.

Метаданные играют важную роль в эффективном управлении и использовании медицинских данных, поддерживая научные исследования, клиническую практику и здравоохранение в целом.

Заключение / Conclusion

Метаданные с каждым годом имеют всё большее влияние на эффективность лечения пациента, преемственность данных между учреждениями и сохранность истории заболевания. Появление таких проектов как MINERVA указывают на потребность системы здравоохранения в новых и качественных системах во всём мире, которые позволят привести данные в упорядоченный вид, и облегчат поиск информации о пациенте.

ДОПОЛНИТЕЛЬНАЯ ИНФОРМАЦИЯ

Конфликт интересов

Авторы декларируют отсутствие конфликтов интересов.

Финансирование

Работа выполнялась без спонсорской поддержки.

Участие авторов

Валеева А. Х.-М., Гомон Ю. М., Белявская Н. М. — написание текста, редактирование, оформление статьи.

СВЕДЕНИЯ ОБ АВТОРАХ

Валеева Алтынай Хаджи-Муратовна — ординатор кафедры клинической фармакологии и доказательной медицины ФГБОУ ВО ПСПбГМУ им. акад. И. П. Павлова Минздрава России, Санкт-Петербург, Российская Федерация

Автор, ответственный за переписку

e-mail: sovoglazka2000@yandex.ru

ORCID ID: 0009-0001-2228-8565

Гомон Юлия Михайловна — д. м. н., профессор кафедры клинической фармакологии и доказательной медицины ФГБОУ ВО ПСПбГМУ им. акад. И. П. Павлова, Санкт-Петербург, Российская Федерация

e-mail: gomonmd@yandex.ru

ORCID ID: 0000-0001-7704-9900

РИНЦ SPIN-код: 1839-9558

ADDITIONAL INFORMATION

Conflict of interests

The authors declare no conflict of interest.

Funding

The work was carried out without sponsorship.

Authors' participation

Valeeva AH-M, Gomon YuM, Belyavskaya NM — text writing, editing, article design.

ABOUT THE AUTHORS

Altynai H.-M. Valeeva — Resident of the department of Clinical Pharmacology and Evidence-Based Medicine, Pavlov First Saint Petersburg State Medical University, St. Petersburg, Russian Federation

Corresponding author

e-mail: sovoglazka2000@yandex.ru

ORCID ID: 0009-0001-2228-8565

Yulia M. Gomon — Dr. Sci. (Med.), professor of the department of Clinical Pharmacology and Evidence-Based Medicine, Pavlov First Saint Petersburg State Medical University, St. Petersburg, Russian Federation

e-mail: gomonmd@yandex.ru

ORCID ID: 0000-0001-7704-9900

RSCI SPIN-code: 1839-9558

Белявская Надежда Максимовна — врач-терапевт ГБУЗ ЛО «Всеволожская клиническая межрайонная больницы» Врачебной амбулатории д. Кальтино, Ленинградская область, Российская Федерация

e-mail: belnad08@gmail.com

ORCID ID: 0009-0000-2215-5861

Nadezhda M. Belyavskaya — Belyavskaya Nadezhda — local doctor therapist of the Vsevolozhsk Clinical Interdistrict Hospital, Kaltino village outpatient clinic, Leningrad Region, Russian Federation

e-mail: belnad08@gmail.com

ORCID ID: 0009-0000-2215-5861

Литература/References

- ГОСТ Р 7.0.10-2010 Набор элементов метаданных «Дублинское ядро» (ISO 15836:2003 Information and documentation – The Dublin Core Metadata Element Set) [GOST R 7.0.10-2010 The Dublin Core Metadata Element Set (ISO 15836:2003 Information and documentation – The Dublin Core Metadata Element Set)].
- Приказ Минздрава РФ от 24.12.2018 № 911Н «Об утверждении требований к государственным информационным системам в сфере здравоохранения субъектов Российской Федерации, медицинским информационным системам медицинских организаций и информационным системам фармацевтических организаций» [Order of the Ministry of Health of the Russian Federation dated December 24, 2018 No. 911n "On Approval of Requirements for State Information systems in the field of healthcare of the Subjects of the Russian Federation, medical information systems of medical organizations and information systems of pharmaceutical organizations"].
- Свердлов Ф. Ю. Проблема информатизации лечебно-профилактических учреждений РФ (на примере ЛПУ г. Москвы). *Врач и информационные технологии*. 2014;4:52-58. [Sverdlov F. Yu. The problem of informatization of medical and preventive institutions of the Russian Federation (on the example of a medical and preventive institution in Moscow). *Doctor and information technology*. 2014;4:52-58. (In Russ.)].
- Good Practice Guide for the use of the Metadata: Catalogue of Real-World Data Sources V 1.0, European Medicines Agency, EMA/787647/2022, September 2022. Доступно по: https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/good-practice-guide-use-metadata-catalogue-real-world-data-sources_en.pdf.
- Gini R, Sturkenboom MCJ, Sultana J, Cave A, Landi A, Pacurariu A, Roberto G, Schink T, Candore G, Slattery J, Trifirò G; Working Group 3 of ENCePP (Inventory of EU data sources and methodological approaches for multisource studies). Different Strategies to Execute Multi-Database Studies for Medicines Surveillance in Real-World Setting: A Reflection on the European Model. *Clin Pharmacol Ther*. 2020 Aug;108(2):228-235. doi: 10.1002/cpt.1833.
- Pajouheshnia R, Gini R, Gutierrez L, Swertz MA, Hyde E, Sturkenboom M, Arana A, Franzoni C, Ehrenstein V, Roberto G, Gil M, Macià MA, Schäfer W, Haug U, Thurin NH, Lassalle R, Droz-Perroteau C, Zaccagnino S, Busto MP, Middelkoop B, Gembert K, Sanchez-Saez F, Rodriguez-Bernal C, Sanfèlix-Gimeno G, Hurtado I, Acosta MB, Poblador-Plou B, Carmona-Pérez J, Gimeno-Miguel A, Prados-Torres A, Schultze A, Jansen E, Herings R, Kuiper J, Locatelli I, Jazbar J, Žerovnik Š, Kos M, Smit S, Lind S, Metspalu A, Simou S, Hedenmalm K, Cochino A, Alcini P, Kurz X, Perez-Gutthann S. Metadata for Data discoverability and Study replicability in observational Studies (MINERVA): Development and Pilot of a Metadata List and Catalogue in Europe. *Pharmacoepidemiol Drug Saf*. 2024 Aug;33(8):e5871. doi: 10.1002/pds.5871.
- Kurz X, Perez-Gutthann S; ENCePP Steering Group. Strengthening standards, transparency, and collaboration to support medicine evaluation: Ten years of the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCEPP). *Pharmacoepidemiol Drug Saf*. 2018 Mar;27(3):245-252. doi: 10.1002/pds.4381.
- Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten JW, da Silva Santos LB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, Gray AJ, Groth P, Goble C, Grethe JS, Heringa J, 't Hoen PA, Hooft R, Kuhn T, Kok R, Kok J, Lusher SJ, Martone ME, Mons A, Packer AL, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone SA, Schultes E, Sengstag T, Slater T, Strawn G, Swertz MA, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data*. 2016 Mar 15;3:160018. doi: 10.1038/sdata.2016.18. Erratum in: *Sci Data*. 2019 Mar 19;6(1):6. doi: 10.1038/s41597-019-0009-6.
- HMA-EMA, "Priority Recommendations of the HMA-EMA Joint big Data Task Force," HMA EMA Big Data Steering Committee Group. Доступно по: <https://www.ema.europa.eu>.

- eu/en/documents/other/priority-recommendations-hma-ema-joint-big-data-task-force_en.pdf.
10. HMA/EMA joint Big Data Steering Group. Big Data Workplan 2023-2025. Доступно по: https://www.ema.europa.eu/en/documents/work-programme/workplan-2023-2025-hma-ema-joint-big-data-steering-group_en.pdf.
11. Pajouheshnia R, Gini R, Gutierrez L, Swertz MA, Hyde E, Sturkenboom M, Arana A, Franzoni C, Ehrenstein V, Roberto G, Gil M, Maciá MA, Schäfer W, Haug U, Thurin NH, Lassalle R, Droz-Perroteau C, Zaccagnino S, Busto MP, Middelkoop B, Gembert K, Sanchez-Saez F, Rodriguez-Bernal C, Sanfélix-Gimeno G, Hurtado I, Acosta MB, Poblador-Plou B, Carmona-Pérez J, Gimeno-Miguel A, Prados-Torres A, Schultze A, Jansen E, Herings R, Kuiper J, Locatelli I, Jazbar J, Žerovnik Š, Kos M, Smit S, Lind S, Metspalu A, Simou S, Hedenmalm K, Cochino A, Alcini P, Kurz X, Perez-Gutthann S. Metadata for Data dIscoverability aNd Study rEplicability in obseRVAtional Studies (MINERVA): Development and Pilot of a Metadata List and Catalogue in Europe. *Pharmacoepidemiol Drug Saf.* 2024 Aug;33(8):e5871. doi: 10.1002/pds.5871.
12. European Medicines Agency. Good Practice Guide for the Use of the Metadata Catalogue of Real-World Data Sources V 1.0. Доступно по: https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/good-practice-guide-use-metadata-catalogue-real-world-data-sources_en.pdf.
13. Swertz MA, Dijkstra M, Adamusiak T, van der Velde JK, Kanterakis A, Roos ET, Lops J, Thorisson GA, Arends D, Byelas G, Muilu J, Brookes AJ, de Brock EO, Jansen RC, Parkinson H. The MOLGENIS toolkit: rapid prototyping of biosoftware at the push of a button. *BMC Bioinformatics.* 2010 Dec 21;11 Suppl 12(Suppl 12):S12. doi: 10.1186/1471-2105-11-S12-S12.
14. European Medicines Agency. “RWD Catalogues: Support. FAQ No. 57”. Режим доступа: <https://catalogues.ema.europa.eu/support>.
15. Gini R, Pajouheshnia R, Gardarsdottir H, Bennett D, Li L, Gulea C, Wientzek-Fleischmann A, Bazelier MT, Burcu M, Dodd C, Durán CE, Kaplan S, Lanes S, Marinier K, Roberto G, Soman K, Zhou X, Platt R, Setoguchi S, Hall GC. Describing diversity of real world data sources in pharmacoepidemiologic studies: The DIVERSE scoping review. *Pharmacoepidemiol Drug Saf.* 2024 May;33(5):e5787. doi: 10.1002/pds.5787.
16. Pajouheshnia R, Gini R, Gutierrez L, Swertz MA, Hyde E, Sturkenboom M, Arana A, Franzoni C, Ehrenstein V, Roberto G, Gil M, Maciá MA, Schäfer W, Haug U, Thurin NH, Lassalle R, Droz-Perroteau C, Zaccagnino S, Busto MP, Middelkoop B, Gembert K, Sanchez-Saez F, Rodriguez-Bernal C, Sanfélix-Gimeno G, Hurtado I, Acosta MB, Poblador-Plou B, Carmona-Pérez J, Gimeno-Miguel A, Prados-Torres A, Schultze A, Jansen E, Herings R, Kuiper J, Locatelli I, Jazbar J, Žerovnik Š, Kos M, Smit S, Lind S, Metspalu A, Simou S, Hedenmalm K, Cochino A, Alcini P, Kurz X, Perez-Gutthann S. Metadata for Data dIscoverability aNd Study rEplicability in obseRVAtional Studies (MINERVA): Development and Pilot of a Metadata List and Catalogue in Europe. *Pharmacoepidemiol Drug Saf.* 2024 Aug;33(8):e5871. doi: 10.1002/pds.5871.
17. Swertz M, van Enckevort E, Oliveira JL, Fortier I, Bergeron J, Thurin NH, Hyde E, Kellmann A, Pajouheshnia R, Sturkenboom M, Cunningham M, Nybo Andersen AM, Marcon Y, Gonçalves G, Gini R. Towards an Interoperable Ecosystem of Research Cohort and Real-world Data Catalogues Enabling Multi-center Studies. *Yearb Med Inform.* 2022 Aug;31(1):262-272. doi: 10.1055/s-0042-1742522.