Journal of Clinical Epidemiology 187 (2025) 111920

KEY CONCEPTS IN CLINICAL EPIDEMIOLOGY

FAIRification of biomedical research data

Ka Hin Tai[a,*], Marcel Müller[b], Ulrich Mansmann[b], Anna Catharina Vieira Armond[c], Evelyne Decullier[d], Anne Le Louarn[e], Nchangwi Syntia Munung[f], Florian Naudet[g,h], Fabian Prasser[i], Ulrich Sax[j]

[a]*Univ Rennes, Inserm, EHESP, Irset (Institut de recherche en santé, environnement et travail) - UMR_S 1085, F-35000, Rennes, France*
[b]*Institute for Medical Information Processing, Biometry, and Epidemiology, Faculty of Medicine, Ludwig-Maximilians-Universität, München, Germany*
[c]*Metaresearch and Open Science Program, University of Ottawa Heart Institute, Ottawa, Canada*
[d]*Hospices Civils de Lyon, Pôle Santé Publique, Unité de Recherche Clinique, Lyon, France*
[e]*GCS CNCR (Comité National de Coordination de la Recherche), Paris, France*
[f]*Division of Human Genetics, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa*
[g]*Université Rennes, CHU Rennes, Inserm, Centre d'Investigation Clinique de Rennes (CIC1414), Service de Pharmacologie Clinique, Institut de Recherche en Santé, Environnement et Travail (Irset), UMR S 1085, EHESP, Rennes, France*
[h]*Institut Universitaire de France (IUF), Paris, France*
[i]*Berlin Institute of Health at Charité, Universitätsmedizin Berlin, Berlin, Germany*
[j]*Department of Medical Informatics, University Medical Center, Göttingen, Germany*

Accepted 31 July 2025; Published online 6 August 2025

Abstract

The Findable, Accessible, Interoperable, and Reusable guiding principles promote Findability, Accessibility, Interoperability, and Reuse of data to enhance data management and stewardship. In biomedicine, particular ethical, legal, and technical barriers complicate research data sharing. To help researchers overcome these challenges, we propose a framework of FAIRification from three dimensions − scientific, technical, and legal/ethical. We advocate for prospective FAIRification of study data, starting with a strong emphasis on planning for data-sharing from the beginning. Reflective questions throughout the process guide researchers to reflect on their situation. Researchers should assess resources and feasibility, secure technical and legal support, consider stakeholder needs, and devise an appropriate data sharing process. Given the sensitivity of biomedical data, confidentiality and security require careful attention. The data sharing strategy should be finalized before the study starts and documented in relevant study materials. Technical preparation for data sharing follows planning. Data should be well-documented with a data dictionary and metadata to facilitate reuse and provided in an accessible format. The data can be hosted on a repository to promote sharing and reuse. While a secure repository provides the technical foundation for data protection, effective administration is required to enforce data use agreements and licensing. We also discuss the importance of subsequent management upon data upload. Continued support for researchers and data maintenance are essential for effective reuse. Examples and resources to facilitate FAIRification are included to help researchers navigate challenges and ensure biomedical data are FAIR and reusable.

Keywords: FAIR principle; Data sharing; Open science; Responsible research; Biomedical research; Metadata standards

Funding: Several authors (K.H., U.M., E.D., F.N., F.P., and U.S.) are members of a doctoral network on data sharing funded by the EU Horizon Marie Skłodowska-Curie Actions (Horizon-MSCA.2022-DN 101120360), which provided a collaborative environment that informed and supported this work. Funded by the European Union. Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the granting authority can be held responsible for them.

* Corresponding author. Pièce 027, Bâtiment 06, Rdc, 01-Villejean, 2 Av. du Professeur Léon Bernard, 35000, Rennes, France.
E-mail address: ka-hin.tai@univ-rennes.fr (K.H. Tai).

https://doi.org/10.1016/j.jclinepi.2025.111920

Data sharing catalyzes clinical research [1] by enabling secondary data use, meta-analyses, validation studies, and informing study planning. The FAIR guiding principles—Findable, Accessible, Interoperable, and Reusable—guide data preparation for sharing, with an emphasis on machine-actionability [2] to maximize data value [3]. FAIRification involves inspecting, describing, preparing, and publishing data [4]. Biomedical data require additional oversight due to their sensitivity.

Despite increasing attention, data sharing rates remain suboptimal [5]. Without clear technical specifications, the FAIR principles remain a moving target. Researchers—especially those with limited resources—often face ethical, legal, and technical barriers to implementation [6]. In this article, we propose a step-by-step FAIRification procedure, guided by reflective questions and tailored to the challenges individual researchers may encounter.

1. FAIRification

Supplementary box 1 defines key terms in the FAIRification process. Before outlining the steps toward FAIRness (Fig), remember to view FAIR from the perspective of the data reuser. Data are only FAIR if it functions at the user end—unfriendly formats will not encourage reuse [7].

While retrospective FAIRification is possible, we strongly encourage prospective planning to avoid unforeseen challenges and costly resource requirements [8]. Early data-sharing preparation enables efficient, reuse-friendly data collection [9]. Embedding FAIR principles and especially (meta)data interoperability from the outset adds flexibility and helps ensure your study remains adaptable and future-proof.

2. Plan

Data sharing is a complex task that requires substantial personnel and technical resources [10]. Involve all key stakeholders early and discuss how your data may be reused. Funders initiate and finance studies; investigators are responsible for data collection; and study participants voluntarily contribute their data—all parties should participate in the decision-making process [11]. If investigators plan for open sharing but the sponsor expects restricted access, this mismatch in expectations can delay or even prevent FAIRification. Data use can be formalized using tools like Data Use Ontology [12] or Open Digital Rights Language [13], which specify usage permissions.

Data Sharing Vision

Who is my target audience?
Do I want to share everything down to individual participant data or just the protocol and blank forms?
Do I share in an open way or with restrictions?
Should the subsequent reuse be independent from my team or collaborative?
Is my data prepared mainly for replication check or inciting new hypothesis?
What are the potential reuse cases?

Your available resources may limit your data sharing options [14]. FAIRification requires professional expertise in managing research data [15,16]. You can either appoint a data steward [17] or contract support from a repository service [18,19]. Consult your institute for IT infrastructure and legal support. Assess feasibility by evaluating your current situation. For example, Inau et al. demonstrated how a self-assessment of a dataset's FAIRness can inform the preparation and sharing of sensitive data [18]. Ethical and legal considerations also determine how and to what extent you can share data. Protecting participant confidentiality is essential [19]. Repository requirements may restrict your choices. Supplementary Table 1 provides an overview of some repository features. If you are designing a study, all data-sharing decisions should be clearly described in a data management and sharing plan (DMSP). Funders (e.g., the National Institutes of Health (NIH) [20]), institutions, and research ethics committees often mandate a DMSP during grant or protocol submission. The data request process should balance feasibility with openness. Share data as openly as possible and as restricted as necessary. While ease of access is ideal, a well-functioning process must always be provided.

Resources, Requirements, and Support

How much time and money can I invest for data sharing?
How much control over the data shared should be retained?
What is the appropriate level of detail for my dataset?
What is requested by the ethical board?
Where can I get technical and legal support?

FAIRification is not just a technical checklist—it draws on specific researcher skills in data handling, documenta-

tion, and legal awareness. As you navigate this process, reflect on the competencies you already have and those you may need to develop. Resources such as Skills4EOSC [21] and FAIRsFAIR [22] projects offer practical guidance to help researchers build digital capabilities and qualify for FAIR and open science.

You and your team are likely the only ones who fully understand your study. Providing your dataset as-is is not only unfair but also risky—external researchers may misinterpret it. It is your responsibility to bring others to your level of understanding regarding the study design, procedures, and data structure [23]. Review your study as if you were an outsider, and ask yourself: what would be unclear to someone unfamiliar with the project? Document the data's provenance so potential users can assess its relevance.

Clarity and Comprehension

What material is missing to help the comprehension of my study?
How much explanation is necessary?
Which part of my study or dataset could be easily misinterpreted?

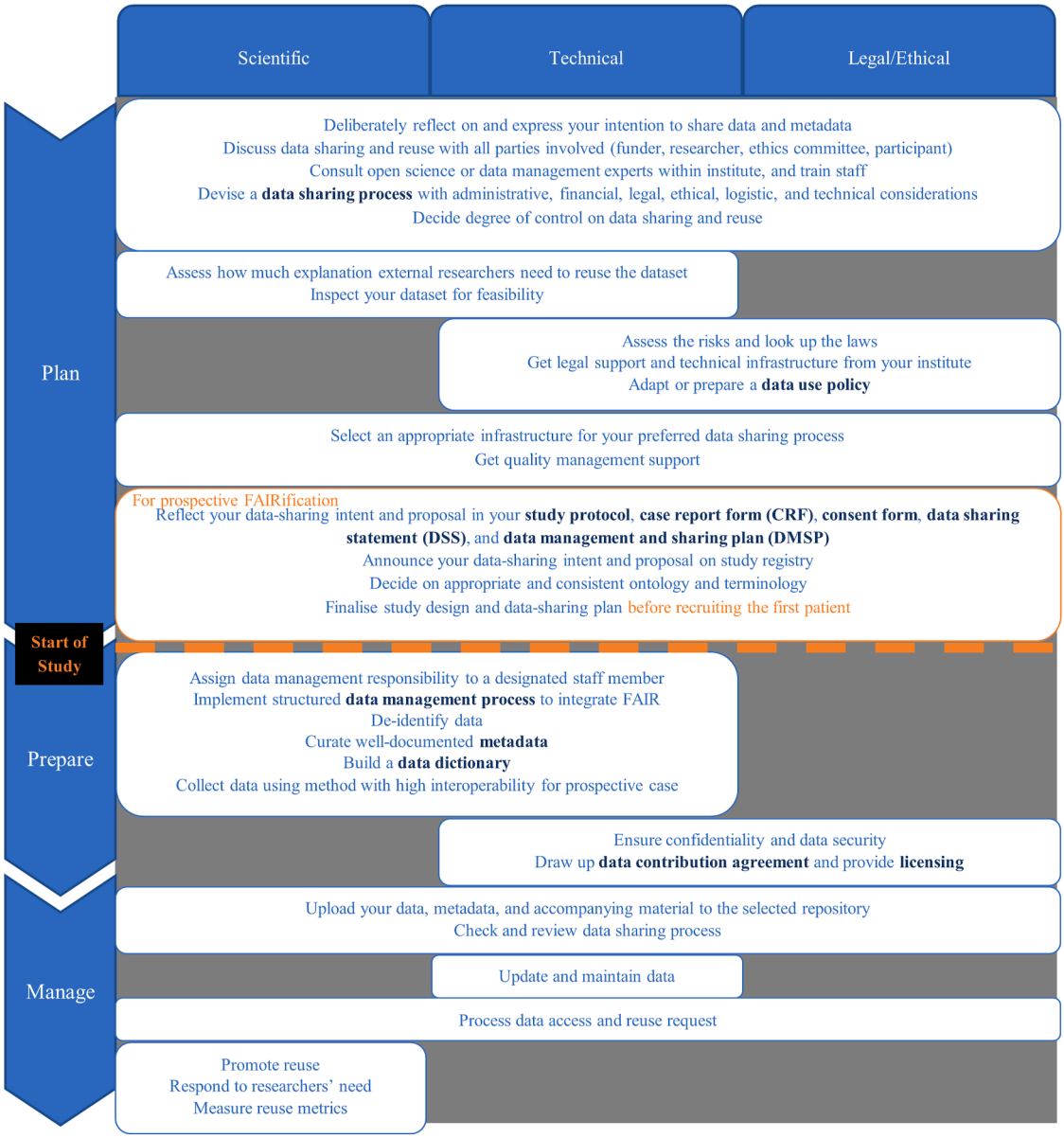

Figure. Dimensions and key outputs of the FAIRification process (outputs of some steps are highlighted in bold. Further details are provided in main text.).

Assess your dataset from a feasibility perspective. These insights can guide your choice of a compatible repository and inform your data sharing plan. Ideally, these steps should be taken in advance. Prospective planning allows greater flexibility—if researchers are aware of potential reuse scenarios, they can design the study and define variables in ways that better support future data reuses.

Data Format, Structure, and Sensitivity

How big is my dataset and in what data format?
Which terminology and ontology are used?
Is my dataset nonproprietary and future-proof?
Is my data thoroughly processed and sealed, or updates are expected from time to time?
How confidential and sensitive is my dataset?
Under what condition have the study participants consented their data to be shared?
What type of variables do I have and are they sufficiently standardized and formatted for analysis?
What analysis tools and dependencies are preferred?

You will need a reliable and widely used data repository to host your dataset. The NIH outlines key criteria for selecting appropriate data repositories [24]—such as long-term sustainability, quality assurance, and clear usage guidance—and maintains a searchable list of suggestions [25]. Supplementary Table 1 presents platforms suitable for individual nonprofit researchers in biomedicine. Stay informed about new repositories relevant to your specialty (eg, omics, imaging, epidemiology). Many repositories serve specific communities and topics, which may impose structural or procedural constraints. For example, OpenNeuro [26] only accepts neuroimaging data organized according to the Brain Imaging Data Structure convention but not others, and NIH Genomic Data Commons [27] does not accept genomic data from participants aged 90 years or older due to privacy and security concerns.

As importantly, you should assess risks related to confidentiality, integrity, and availability before sharing your data. Elements such as Safe Projects, Safe People, and Safe Settings from the Five Safes framework [28] often depend on repository infrastructure. Once risks are identified, propose appropriate mitigation strategies. While ethics committees should oversee this process, success ultimately requires team effort. Always follow applicable journal policies, institutional guidelines, and regional and international regulations—especially for cross-border data sharing [29].

Data Security and Risk

What are the possible scenarios of data misuse?
How secure is my data in the repository selected and in my data sharing process?
Is there a potential harm to participants in any way, including reidentification of patients?
Are there scientific harms in terms of poor reuse, data misuse due to lack of documentation?

The intent to share data should be transparent from the study design phase and clearly reflected in related documents and forms. Participant consent for data sharing must be obtained early during recruitment and data collection, with explicit information about what will be shared and how [30], and options allowing participants to withdraw their consent later. All data-sharing plans should be finalized before the first participant is enrolled. FAIRification should be integral to the study workflow—not an afterthought.

3. Prepare

Data preparation should start as early as possible. Here we follow the ten-step workflow from Sinaci et al.[6] and highlight specific considerations. Biomedical research often involves sensitive personal data; therefore, deidentification must be prioritized to minimize privacy and confidentiality risks [31]. Many open tools are available for pseudonymization and anonymization [32]. Timestamp variables should be accompanied by, or even replaced with, relative duration variables, for example, date of examination becoming days since patient recruitment. At the study design stage, ensure your data collection methods integrate FAIR principles. For example, using structured instruments and standardized variables such as SNOMED CT (Systematized Nomenclature of Medicine - Clinical Terms) and LOINC (Logical Observation Identifiers Names and Codes) code [33] can improve interoperability. Ultimately, reusability depends on the intrinsic quality of data and how rigorously it was collected [34]—datasets with inconsistent formats, missing key variables, or poorly defined measures are unlikely to support meaningful secondary analysis.

Data Quality and Reusability

Is my data collection process consistent and reproducible?
Are my variables standardized and coded using recognized vocabularies?
What factors in my dataset could reduce its reusability (eg, missing data, inconsistencies)?

To determine whether they can work with your dataset, other researchers need contextual information about your study. Metadata provide this context and explain how the data can be used. For meta-analysts, metadata are especially important to assess study heterogeneity [35]. When indexed on repositories, metadata also improve discoverability. Follow established schemas such as the CDISC Study Data Tabulation Model (SDTM) [36] or the OpenAIRE Guidelines [37,38]. A data dictionary should accompany all analyzable variables, detailing validation rules, sampling methods, format, and limitations. It should enable users to clearly distinguish between variables. Precision and context are key—measurement units, coded options, and the exact timing or conditions of data collection all make the dataset more reusable. In prospective FAIRification, metadata and data dictionaries can be developed as early as the protocol stage.

Metadata and Documentation

Is the metadata clear and comprehensive enough to explain my study and dataset?
Is the data dictionary understandable and unambiguous, so it will not confuse external statisticians?

Your data sharing process should follow well-defined, standardized procedures documented in standard operating procedures (SOPs). SOPs promote consistency, transparency, and alignment with best practices, thereby improving overall data management quality. They also serve as valuable resources for successors, supporting smooth onboarding while ensuring continuity and reliability in data sharing practices.

Internal Processes and SOPs

Are the SOPs accessible, understandable, and version-controlled?
How well would a new team member understand how to share data by reading our documentation?
Have I tested the SOPs in practice?

Beyond the workflow, data security must also be assured. While you may have limited control over repository infrastructure, data security is only as good as the platform you choose. Assess the repository's trustworthiness (e.g., refer to Supplementary Table 1) and sign a data distribution agreement reviewed by legal experts and data protection officers. Verify participant consent and remove records from those who no longer agree to share. Prepare key documents—including the DMSP and Data Sharing Statement (DSS). The DMSP should clearly describe procedures from collection to sharing. Since 2017, the International Committee of Medical Journal Editors-affiliated journals require a DSS as a condition for clinical trial publication [39], which indicates the intention, method, and context of data sharing in the presented clinical trial. These documents should be attached to both the repository and study registry (e.g., Clinicaltrials.gov).

4. Manage

Voilà, once your data are ready, it is eventually time to upload them to the selected repository. Lock your dataset, and if changes are needed, upload them as a new version. Thoroughly review the platform's policy to ensure your data sharing aligns with relevant legal, ethical, and institutional guidelines. Test whether your data request process functions as intended, and assess data reusability through independent use [40]. A digital object identifier will provide your dataset with a permanent and unique link for identification, access, and citation. You can, for example, promote reuse by linking the dataset to your Open Researcher and Contributor ID profile.

Data Access and Linking

How does the data access process work exactly for an external user?
Is my dataset clear linked to supporting documents (protocol, metadata)?

Once your dataset is publicly available, several tasks require ongoing attention. These include data maintenance during system or content updates, handling data access requests and legal agreements, and responding to researchers' feedback—others may suggest more practical ways to work with your data. Document and track data reuse and its impact for transparency. You can start by recording the number of data access requests received as a basic metric of outreach and interest, and build on that with citations, acknowledgments, or documented collaborations. A technical and administrative maintenance plan should be outlined in the SOPs to ensure continuity and accountability.

Maintenance and Monitoring

What updates may occur in the future (eg, data correction, versioning)?
Who in my team is responsible for postupload maintenance?
How will I monitor and respond to data reuse?

FAIR is a strong foundation for enabling data sharing. It is, however, not sufficient for achieving meaningful reuse [5,7]. Data quality, resource investment, community

engagement, and social acceptance of reuse are equally important. Academic institutions should lead by example—mandating data sharing, providing adequate infrastructure and resources [10], developing practical guidelines, and offering training and audit services. The research community must also reshape incentives to reward open science [41]. The true value of FAIR data lies in its reuse. While you are learning how to FAIRify your data, do not forget to capitalize on the invaluable assets available — by reusing data shared by other researchers to drive further discoveries and consolidate existing knowledge!

CRediT authorship contribution statement

Ka Hin Tai: Writing — review & editing, Writing — original draft, Visualization, Project administration, Investigation, Data curation, Conceptualization. **Marcel Müller:** Writing — review & editing, Writing — original draft, Investigation, Data curation, Conceptualization. **Ulrich Mansmann:** Writing — review & editing, Writing — original draft, Investigation, Data curation, Conceptualization. **Anna Catharina Vieira Armond:** Writing — review & editing, Writing — original draft, Investigation, Data curation, Conceptualization. **Evelyne Decullier:** Writing — review & editing, Writing — original draft, Investigation, Data curation, Conceptualization. **Anne Le Louarn:** Writing — review & editing, Writing — original draft, Investigation, Data curation, Conceptualization. **Nchangwi Syntia Munung:** Writing — review & editing, Writing — original draft, Investigation, Data curation, Conceptualization. **Florian Naudet:** Writing — review & editing, Writing — original draft, Supervision, Project administration, Investigation, Data curation, Conceptualization. **Fabian Prasser:** Writing — review & editing, Writing — original draft, Investigation, Data curation, Conceptualization. **Ulrich Sax:** Writing — review & editing, Writing — original draft, Supervision, Project administration, Investigation, Data curation, Conceptualization.

Declaration of competing interest

K.H.T., F.N., U.M., F.P., E.D., U.S. reports administrative support, travel, and writing assistance were provided by SHARE-CTD doctoral network (Horizon-MSCA.2022-DN 101120360). U.S. reports a relationship with i2b2 tranSMART foundation that includes: board membership; with NFDI4health that includes: funding grants, nonfinancial support, and travel reimbursement; with German Research Foundation that includes: travel reimbursement; and with German Ministry of Education and Research that includes: travel reimbursement. F.N. reports a relationship with French National Research Agency that includes: funding grants; with French Ministry of Health that includes: funding grants; and with French Ministry of Research that includes: funding grants. N.S.M. reports a relationship with National Institute of Mental Health that includes: funding grants and with National Heart Lung and Blood Institute that includes: funding grants. There are no competing interests for any other author.

Acknowledgments

We acknowledge Dr. Marie Gebhardt and Karoline Buckow for their valuable contributions.

Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.jclinepi.2025.111920.

Further readings

National Institutes of Health. Data Management and Sharing Policy | Data Sharing [Internet]. sharing.nih.gov. Available from: https://sharing.nih.gov/data-management-and-sharing-policy - A comprehensive guideline covering responsible data sharing, privacy protection, data management, and budgeting.

Pellen C, Anne Le Louarn, Gilliosa Spurrier-Bernard, Decullier E, Jean-Marie Chrétien, Rosenthal E, et al. Ten (not so) simple rules for clinical trial data-sharing. PLOS Computational Biology. 2023 Mar 9;19(3):e1010879—9. — 10 considerations facilitating clinical trial data-sharing, such as funding and data protection requirement.

Tudur Smith C, Nevitt S, Appelbe D, Appleton R, Dixon P, Harrison J, et al. Resource implications of preparing individual participant data from a clinical trial to share with external researchers. Trials. 2017 Jul 17;18(1). — Example of retrospective FAIRification of clinical trial data with details on anonymisation process, data pack preparation, division of labor, and resource implication in time and cost.

J. Elis Hoffmann, Hanß S, Kraus M, Schaller J, Schäfer C, Stahl D, et al. The DZHK research platform: maximization of scientific value by enabling access to health data and biological samples collected in cardiovascular clinical studies. Clinical Research in Cardiology. 2023 Mar 8;112 (7):923—41. — Example of establishment of data sharing platform with details on data standardization, storage, ethical consideration, and use and access policy.

Robert Andrews, Andrew Mason, Sara Morsy, Philippe Rocca-Serra, Xenia Perez Sitja, Branka Franicevic, Katarzyna Kamieniecka, Khaled Jum'ah, Krzysztof Poterlowicz, FAIRification of an RNAseq dataset (Galaxy Training Materials). https://training.galaxyproject.org/training-material/topics/fair/tutorials/fair-rna/tutorial.html Online; accessed Mon Feb 24 2025 — Example of FAIRification of RNAseq dataset with emphasis on the putting each point of the FAIR principles into practice.

Data availability

No data was used for the research described in the article.

References

[1] Mansmann U, Locher C, Prasser F, Weissgerber T, Sax U, Posch M, et al. Implementing clinical trial data sharing requires training a new generation of biomedical researchers. Nat Med 2023;29(2):298−301.

[2] Wilkinson MD, Dumontier M, Aalbersberg IjJ, Appleton G, Axton M, Baak A, et al. The FAIR guiding principles for scientific data management and stewardship. Sci Data 2016;3(1):160018.

[3] Mons B, Neylon C, Velterop J, Dumontier M, da Silva Santos LOB, Wilkinson MD. Cloudy, increasingly FAIR; revisiting the FAIR data guiding principles for the European open science cloud. Inf Serv Use 2017;37(1):49−56.

[4] GO FAIR. FAIRification process. GO FAIR. Available at: https://www.go-fair.org/fair-principles/fairification-process/. Accessed April 1, 2025.

[5] Ohmann C, Moher D, Siebert M, Motschall E, Naudet F. Status, use and impact of sharing individual participant data from clinical trials: a scoping review. BMJ Open 2021;11(8):e049228.

[6] Sinaci AA, Núñez-Benjumea FJ, Gencturk M, Jauer ML, Deserno T, Chronaki C, et al. From raw data to FAIR data: the FAIRification workflow for health research. Methods Inf Med 2020;59(S 01): e21−32.

[7] Locher C, Le Goff G, Le Louarn A, Mansmann U, Naudet F. Making data sharing the norm in medical research. BMJ 2023;382:p1434.

[8] Tudur Smith C, Nevitt S, Appelbe D, Appleton R, Dixon P, Harrison J, et al. Resource implications of preparing individual participant data from a clinical trial to share with external researchers. Trials 2017;18(1):319.

[9] Jorgenson LA, Wolinetz CD, Collins FS. Incentivizing a new culture of data stewardship. JAMA 2021;326(22):2259.

[10] Borgman CL, Bourne PE. Why it takes a village to manage and share data. arXiv.org. 2021. Available at: https://arxiv.org/abs/2109.01694. Accessed January 6, 2025.

[11] Pellen C, Anne LL, Spurrier-Bernard G, Decullier E, Chrétien J-M, Rosenthal E, et al. Ten (not so) simple rules for clinical trial data-sharing. PLOS Comput Biol 2023;19(3):e1010879.

[12] EMBL-EBI. Ontology lookup service (OLS). Ebi.ac.uk. 2025. Available at: https://www.ebi.ac.uk/ols4/ontologies/duo?viewMode=list. Accessed February 24, 2025.

[13] Pandit HJ, Esteves B. Enhancing data use ontology (DUO) for health-data sharing by extending it with ODRL and DPV. Semantic Web 2024;14:1−26.

[14] Leonelli S. Open data: curation is under-resourced. Nature 2016;538 (7623):41.

[15] Centre for Journalology. Centre for Journalology. Centre for Journalology. 2022. Available at: https://ohri.ca/journalology/ottawa-data-champions. Accessed January 6, 2025.

[16] Higman R. Data champions | research data management. Cam.ac.uk. 2016. Available at: https://www.data.cam.ac.uk/intro-data-champions. Accessed January 6, 2025.

[17] Dumontier M, Wesley K. Advancing discovery science with FAIR data stewardship: findable, accessible, interoperable, reusable. Serials Librarian 2018;74(1−4):39−48.

[18] Inau ET, Dedié A, Anastasova I, Schick R, Zdravomyslov Y, Fröhlich B, et al. The journey to a FAIR CORE DATA SET for diabetes research in Germany. Sci Data 2024;11(1):1159.

[19] Mello MM, Lieou V, Goodman SN. Clinical trial participants' views of the risks and benefits of data sharing. New Engl J Med 2018;378 (23):2202−11.

[20] National Institutes of Health. Data management and sharing policy | data sharing [Internet]. sharing.nih.gov. Available at: https://sharing.nih.gov/data-management-and-sharing-policy. Accessed April 1, 2025.

[21] GARR C. Skills4EOSC [internet]. Skills4eosc. 2024. Available at: https://www.skills4eosc.eu/. Accessed April 1, 2025.

[22] FAIRsFAIR "Fostering FAIR Data Practices In Europe". FAIRsFAIR. Fairsfair.eu. 2021. Available at: https://www.fairsfair.eu/. Accessed June 16, 2025.

[23] Longo DL, Drazen JM. Data sharing. New Engl J Med 2016;374(3): 276−7.

[24] National Institutes of Health. Selecting a data repository | data sharing [Internet]. sharing.nih.gov. Available at: https://sharing.nih.gov/data-management-and-sharing-policy/sharing-scientific-data/selecting-a-data-repository. Accessed April 1, 2025.

[25] National Institutes of Health. Repositories for sharing scientific data | data sharing [internet]. Sharing.nih.gov. Available at: https://sharing.nih.gov/data-management-and-sharing-policy/sharing-scientific-data/repositories-for-sharing-scientific-data. Accessed April 1, 2025.

[26] Stanford center for reproducible neuroscience. OpenNeuro [internet]. openneuro.org. 2019. Available at: https://openneuro.org/. Accessed April 1, 2025.

[27] National Institutes of Health. Home | NCI genomic data commons [Internet]. gdc.cancer.gov. Available at: https://gdc.cancer.gov/. Accessed April 1, 2025.

[28] Desai T, Ritchie F, Welpton R. Five Safes: designing data access for research [Internet]. Bristol: University of the West of England, Bristol Business School, Department of Accounting, Economics and Finance; 2016: (Economics Working Paper Series; 20161601). Available at: https://ideas.repec.org/p/uwe/wpaper/20161601.html. Accessed April 1, 2025.

[29] van de Wiel H, Burgwinkel C, Held L. Sharing clinical trial data [internet]. Zenodo. 2024. Available at: https://zenodo.org/records/13860164. Accessed October 5, 2024.

[30] Ohmann C, Banzi R, Canham S, Battaglia S, Matei M, Ariyo C, et al. Sharing and reuse of individual participant data from clinical trials: principles and recommendations. BMJ Open 2017;7(12): e018647.

[31] Hughes S, Wells K, McSorley P, Freeman A. Preparing individual patient data from clinical trials for sharing: the GlaxoSmithKline approach. Pharm Stat 2014;13(3):179−83.

[32] Haber AC, Sax U, Prasser F, NFDI4Health Consortium. Open tools for quantitative anonymization of tabular phenotype data: literature review. Brief Bioinformatics 2022;23(6):1−10. https://doi.org/10.1093/bib/bbac440.

[33] Dugas M. Portal of medical data models (MDM-Portal) [internet]. Uni-heidelberg.de. 2025. Available at: https://mdm.mi.uni-heidelberg.de/. Accessed February 24, 2025.

[34] Weissgerber TL, Gazda MA, Gustav N, Gerben ter R, Cobey KD, Prieß-Buchheit J, et al. Understanding the provenance and quality of methods is essential for responsible reuse of FAIR data. Nat Med 2024;30(5):1220−1.

[35] Hopkins AM, Modi ND, Abuhelwa AY, Kichenadasse G, Kuderer NM, Lyman GH, et al. Heterogeneity and utility of pharmaceutical company sharing of individual-participant data packages. JAMA Oncol 2023;9(12):1621−6.

[36] Clinical Data Interchange Standards Consortium. SDTM metadata submission guidelines v2.0 [internet]. Cdisc.org. 2021. Available at: https://www.cdisc.org/standards/foundational/sdtm/sdtm-metadata-submission-guidelines-v2-0. Accessed January 9, 2025.

[37] OpenAIRE. OpenAIRE guidelines for literature repositories v3 — OpenAIRE guidelines documentation [Internet]. Openaire.eu. 2020. Available at: https://guidelines.openaire.eu/en/latest/literature/index_guidelines-lit_v3.htm. Accessed January 9, 2025.

[38] Hughes LD, Tsueng G, DiGiovanna J, Horvath TD, Rasmussen LV, Savidge TC, et al. Addressing barriers in FAIR data practices for biomedical data. Sci Data 2023;10(1):98.

[39] Taichman DB, Sahni P, Pinborg A, Peiperl L, Laine C, James A, et al. Data sharing statements for clinical trials: a requirement of the international committee of medical journal. Ann Intern Med 2017;167(1):63−5.

[40] Naudet F, Sakarovitch C, Janiaud P, Cristea I, Fanelli D, Moher D, et al. Data sharing and reanalysis of randomized controlled trials in leading biomedical journals with a full data sharing policy: survey of studies published in the BMJ and PLOS medicine. BMJ 2018;360:k400.

[41] Moher D, Bouter L, Kleinert S, Glasziou P, Sham MH, Barbour V, et al. The Hong Kong principles for assessing researchers: fostering research integrity. PLoS Biol 2020;18(7):e3000737.