Ref. Ares(2024)9005321 - 16/12/2024

ERA4Health Partnership

Deliverable 16.5

Guidelines for data sharing of investigator-initiated clinical studies

WP16

Co-funded by the European Union

Funded by the European Union under the Horizon Europe Framework Programme. Grant Agreement Nº: 101095426.

Technical References

Deliverable No.	**D16.5**
Dissemination Level[1]	PU
Work Package	WP16
Lead beneficiary	ECRIN
Version	1
Due date of deliverable	30/11/2024
Actual submission date	30/11/2024

Versions

Version	Person	Partner	Date
#1.	Salma Malik, Marta del Álamo, Sabrina Lémeret, Mihaela Matei, Sergio Contrino, Jacques Demotes, Sebastian Klammt	ECRIN	18 October 2024

Approved by Coordinator on: 11/11/2024

Disclaimer. Funded by the European Union under the Horizon Europe Framework Programme. Grant Agreement №: 101095426. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Health and Digital Executive Agency (HADEA). Neither the European Union nor the granting authority can be held responsible for them

1 PU = Public
SEN = Sensitive

D16.5. GUIDELINES FOR DATA SHARING OF INVESTIGATOR-INITIATED CLINICAL STUDIES

TABLE OF CONTENTS

GLOSSARY / LIST OF ACRONYMS

Including: a brief definition of the terminology included in this deliverable that allows the reader to understand better the scope of it; list of acronyms.

ANR	Agence Nationale De La Recherche (France)
ASRT	Academy Of Scientific Research And Technology
CERTH	Centre For Research And Technology Hellas (Greece)
DOI	Digital Object Identifier
DPIA	Data Protection Impact Assessment
ECRIN	European Clinical Research Infrastructure Network
ERA4Health	European Research Area For Health
EU	European Union
FAIR	Findable, Accessible, Interoperable And Reusable
GDPR	General Data Protection Regulation
HDABs	Health Data Access Bodies
HIPAA	Health Insurance Portability and Accountability Act
IPD	Individual Patient Data
HRB	The health research board (Ireland)
ICF	Informed Consent Form
ISCIII	Instituto De Salud Carlos III (Spain)
ISO	International Organization for Standardization
LCS	Latvijas Zinatnes Padome (Latvia)
NIH	The National Institutes Of Health (UK)
RCN	Norges Forskningsrad (Norway)
SDTM	Study Data Tabulation Model
VACCELERATE	European Corona Vaccine Trial Accelerator Platform
YODA	The Yale University Open Data Access
ZonMw	Zorgonderzoek Nederland Zon (Netherlands)
HL7	Health Level 7 International
NCDPD	National Council for Prescription Drug Programs
IHTSDO	International Health Terminology Standards Development Organizations
CDISC	Clinical Data Interchange Standards Consortium

EXECUTIVE SUMMARY

The objective of this task is to provide general guidance for preparing data sharing plan for clinical studies, including guidance on informed consent (ICF) for secondary use of data and long-term storage of data, especially Individual Patient Data (IPD) in repositories following the FAIR (Findable, Accessible, Interoperable and Reusable) principles. This guide outlines the process of developing a data sharing plan GDPR compliant and in alignment with the EU and international funders expectations and provides practical guidance on how to meet them. It should be of relevance to stakeholders (funders, grant applicants, coordinating-investigators, research staff, patients' groups, researchers, academia, professional groups, industry, reviewers, and regulatory and ethics authorities) involved in investigator initiated multinational clinical trials.

1. PURPOSE AND OBJECTIVES

A fundamental issue for funders involved in clinical research is to ensure that the science they fund meets the highest research integrity standards. As part of the clinical research activities, clinical trials generate one of the highest levels of evidence of clinical research and therefore the data generated should benefit the whole research community by being available for reuse (1). Indeed, it has been argued that clinical trial data should be shared and treated as a public good whoever generates it, that is whether they are created by publicly funded or by commercial research. Data-sharing aims to foster the benefits that can arise from the use of individual patient data (IPD) by exploring new or unresolved issues from completed trials, by merging them in large IPD meta-analyses or by re-analysing the initial trial data (2). It also guarantees transparency, openness, and reproducibility and honours the generosity and the risk taken by clinical trial participants increasing the value of their contribution (2)(3)(4). Moreover, the artificial intelligence era boosts the value of individual patient level data sharing, as a prerequisite to train algorithms having the potential to improve healthcare strategies. Therefore, easy access to high quality data may also contribute to technological innovation.

The WHO made a Joint Statement on public disclosure of results from clinical trials (5), and various calls to action from clinical research stakeholders to data generators, such as pharmaceutical companies, universities, charities, regulatory agencies, have led to the implementation of policies and recommendations to responsibly share clinical trial data (6). Data-sharing platforms such as Clinical Study Data Request (7), the Yale University Open Data Access (YODA) (8) Project and Vivli (9) were created to facilitate data-sharing from clinical studies and to ensure that clinical trial data are FAIR (10). The International Committee of Medical Journal Editors (ICMJE) has promoted data sharing by requiring authors to include a data-sharing statement in published articles and to register a data-sharing plan for any new trial (11). Nevertheless, to be effective, these initiatives need to be supported by funders, as implementation of data sharing plans requires dedicated financial resources and involves considerable challenges, including maintenance of privacy and confidentiality and patient consent (12).

Several funders (e.g., NIH in the USA, Wellcome Trust, Bill and Melinda Gates Foundation and the European Commission) require a data management plan that includes provision for data sharing, submitted along with the grant application and some established dedicated data repositories. The MRC, UKCRC, Cancer Research UK, and Wellcome Trust have endorsed good practice principles for sharing IPD from publicly funded clinical trials (13). However, these are only recommendations and are not obligatory. Evidence suggests that IPD data sharing does not regularly occur (14).

Funders expect data plans to outline how data will be created, managed, shared, and preserved, explaining any restraints that ought to be employed. Data sharing plans are an opportunity for the grant applicants to demonstrate their perception of good practice and reassure funders that their proposal is in line with their data polices. Funders are therefore key players in supporting data-sharing and are expected

to provide appropriate guidance and to require best practice from their grant recipients (15). Not all funders offer detailed guidance for researchers to develop their plans.

The critical question now is how the existing knowledge can be used by funders to guide the researchers to address the numerous global challenges associated with preparing data sharing plans that should be in alignment with their expectations. To achieve this objective, ERA4Health task 16.6 has worked on providing guidelines to facilitate grant applicants on how to prepare a data sharing plan in alignment with the EU and international funders expectations and provides practical guidance on how to meet these potentials. These guidelines will be included into the ERA4Health upcoming call text for clinical studies.

2. SCOPE

These guidelines should be of relevance to stakeholders (funders, grant applicants, co-investigators, research staff, patients' groups, researchers, academia, professional groups, industry, reviewers, and regulatory authorities and ethics committees) involved in clinical studies with special focus on non-commercial/investigator-initiated trials in Europe. Throughout this document we use IPD (individual patient data) to refer to *all* the participant data available from a trial, and not just the data supporting the conclusions of a specific published paper. Such data will therefore be the datasets used for the various analyses, after appropriate pseudonymisation or anonymisation measures have been applied. Besides the IPD sets, other clinical trial data sources should be made available for sharing (e.g., protocols, clinical study reports, statistical analysis plans, blank consent forms) to enable a full understanding of any data set.

3. METHODOLOGY

These guidelines (Figure 1) are drafted for preparing a data sharing plan in accordance with principles and recommendations for data sharing and reuse of individual participant data (IPD) from clinical trials, provided by Ohmann et al (16) and Kuntz et al (17). Additionally, these guidelines are based on the practical experience of implementation of these strategies and recommendations for promoting data sharing and future reuse in the EU-funded VACCELERATE project (18). Many legal and regulatory texts govern data-sharing and reuse, especially in terms of data protection. Among the strongest requirements, the European General Data Protection Regulation applies to any organization established in the territory of the European Union or processing data from people residing there. Among other requirements, it imposes the design and maintenance of specific records (e.g., records of processing activities); the provision of information about the reuse is mandatory and the processing of health data is prohibited without a specific exemption, for example for research. While the regulation seeks to safeguard the privacy of the individual, there are differing standards and protections. Indeed, within Europe, there has been a fragmented approach to the implementation of the GDPR at a national level. This is a challenge for research collaboration and the sharing of IPD (19) and becomes even more problematic for the cross-

border sharing of data. Since regulatory requirements differ across the world and no harmonization yet exists, it is important to consider the regulatory context of both the data generator and the re-user.

In spring 2024, the European Parliament and the Council reached a political agreement on the Commission proposal for the European Health Data Space (EHDS) Regulation. The EHDS Regulation addresses both the primary use and the secondary use of health records. As the definition of Electronic Health Records (EHRs) includes not only health records, pathogens, health claims and reimbursements, genetic data, public health registry information, wellness data and information on healthcare resources, expenditure and financing, but also covers cohort and clinical trial data, clinical research data sharing is governed by the provisions for secondary use of EHRs in the EHDS Regulation. For a purpose of research, innovation, policy-making, education and patient safety, the EHDS Regulation describes the process for making data accessible, either as downloadable data if fully anonymised, or accessible in a secure processing environment for pseudonymised data (according to the GDPR, anonymization requires ruling out any risk of re-identification). Requests to access the data for secondary use will be assessed by national (or regional) Health Data Access Bodies (HDABs), in charge of delivering a data permit specifying the conditions for accessing the data (nature of data minimization based on the objective of the secondary use, possibility to anonymise the requested data, or need to keep the dataset as pseudonymised and to provide access in the secure processing environment). With regards to informed consent, the EHDS Regulation secured the right for patients to opt-out of the secondary use (with certain exceptions for public-interest, policy-making or statistics purposes, and protections for intellectual property rights and trade secrets when relevant data is shared for secondary use). However, the default option is the possibility for a secondary use of data without collection of an informed consent. But this opt-out provision only applies if the data permit is provided by a HDAB. Instead of using a HDAB, an alternative solution is to use a research infrastructure with ERIC status (e.g. ECRIN-*ERIC*) as a research access body to a data sharing repository, but in such case the opt-out is not valid and the collection of informed consent for data sharing has to be considered (at least in some countries, depending on the GDPR implementation). The EHDS Regulation will be progressively implemented from January 2025 on, and full implementation will not happen before 2029.

National data protection authorities will monitor the enforcement of health data access permits and will be empowered to issue fines in the event of shortcomings.

This complex regulatory environment reinforces the need for good communication between data generators and data re-users in order to align on the essential information about the requirements in force and the procedures that are to be followed. Thus, researchers should seek support from their institutions and approach data teams if they are identifiable.

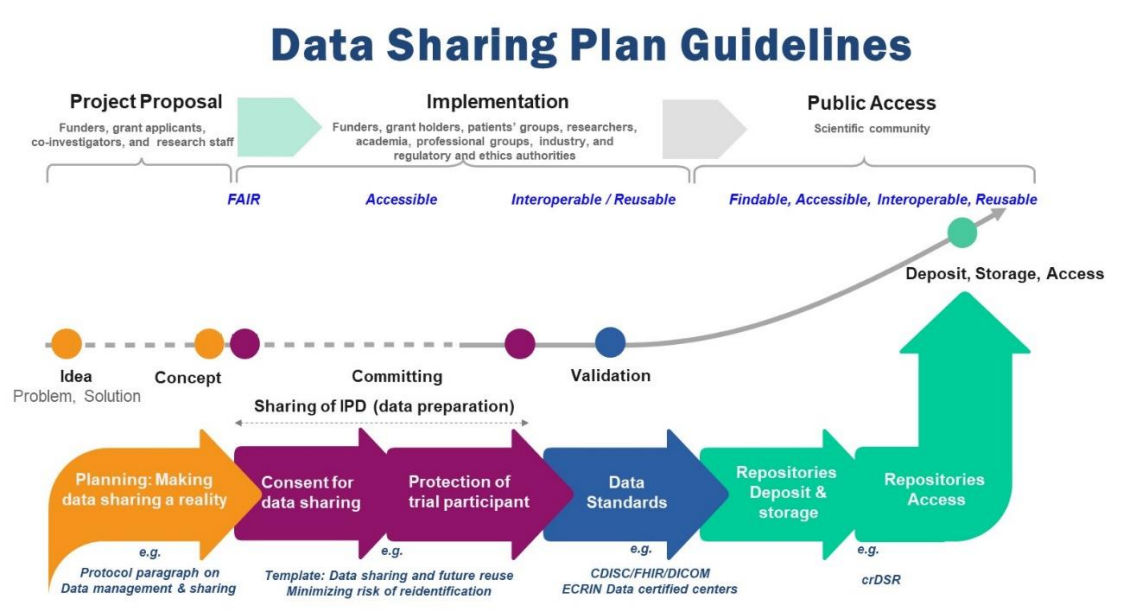


Figure 1: Data sharing plan guidelines

4. DATA SHARING PLAN PREPARATION GUIDELINES

The guidance presented here is structured in correspondence to figure 1. These guidelines are meant to promote and support data sharing and reuse among researchers, adequately inform trial participants and protect their rights and provide effective and efficient systems for preparing, storing and accessing data. By enforcing implementation of this guidelines, funders of clinical trials contribute to the principle that publicly funded research data are a public good, produced in the public interest and should be made openly available with as few restrictions as possible in a timely and responsible manner (20). Next to each of the subsections you will find a summary of the topic to be covered in the data sharing plan along with the useful resources that will facilitate to prepare high quality data sharing plan.

Points to consider	Information to be provided in the plan	Guidance
PLANNING		
Making data sharing a reality **Describe research data management and sharing (Data Management Plan)**	Researchers should ensure that data sharing is considered from the very beginning of study planning and should be included within the trial protocols, data management, sharing plan and trial´s registry entry as well as in patients/participants information and informed consent, if appropriate (see below). Funds for responsible data sharing should be requested by the original trial team from trial funders as part of initial trial grant applications, e.g. to fund dataset preparation and anonymisation/ pseudonymisation Reasonable costs may be recovered from data requesters if appropriate (as mentioned in the EHDS Regulation), but data sharing activities should not be profit-generating	A specific paragraph within a designated section of the protocol should be included. The protocol of the trial should clearly indicate that the data, study documents and participant-level dataset (either as anonymised or pseudonymised), will be made available and summarised in the relevant section of the trial registration record. Recommended resources: • ***SPIRIT 2013 Statement****: Defining Standard Protocol Items for Clinical Trials, item 31a*(21) • ***Annex 1****: **Example of data sharing specific paragraph to be included in the clinical study protocol*** • ***Data management and sharing plan templates for consideration*** https://dmponline.dcc.ac.uk/public_plans • ***International trial´s registry for consideration*** https://clinicaltrials.gov/ • ***Primary registries in the WHO registry network for consideration*** https://www.who.int/clinical-trials-registry-platform/network/primary-registries
IMPLEMENTATION (data collection and preparation)		
Consent for data sharing	All the trial documents (e.g., participants' information leaflet and consent forms, ethical submission documents) should be written and translated taking into account the planned data sharing strategy.	IPD sharing should be based on explicit broad consent by trial participants (or if applicable by their legal representatives) to the sharing and reuse of their data for scientific purposes. The process of informing trial participants about possible sharing of their data, and then gaining their explicit consent to it, is of fundamental importance, and is normally a prerequisite for the sharing of pseudonymised data. The consent for secondary use of IPD should be as broad as

		possible. No consent is needed if the data are to be shared only after anonymisation, or according to the opt-out procedure after obtention of a data permit from a HDAB, or because of the 'public interest' underpinning the GDPR implementation in some countries. ***Annex 2***: In the context of the VACCELERATE project (Task 9.3 Promotion of activities data sharing and future reuse), an informed consent including clauses for data sharing and reuse for future research purposes was generated.
Protection of trial participants **GDPR/Anonymisation/Pseudonymisation**	Shared IPD from clinical trials used for further scientific research should be stored in a secure processing environment and, depending on the request for secondary use, will be made accessible either as pseudonymised or anonymised datasets. Prepare the dataset for sharing (pseudonymisation, minimisation, anonymisation if relevant), taking into account the original patient consent and method of data transfer. Dataset preparation should be done by individuals with an understanding of data management and basic statistics, with quality control provided by a further individual who is independent of the process	Sharing of pseudonymous data should be considered as the default option – this means making data accessible to re-users, after obtention of a data permit, in a secure processing environment, without the possibility to download the data. Depending on the objective of the request for secondary use, the data should be minimised, i.e. removal of data unnecessary for the planed secondary use. Depending on the request for secondary use, sharing anonymised data may be possible if the anonymisation process does not alter the informative value of the dataset. Anonymisation is based on the risk of re-identification, and anonymised data is not governed by the GDPR, therefore anonymised data can be downloaded and circulated. Various anonymisation techniques can be used by a data expert or through anonymization tools like Amnesia https://amnesia.openaire.eu/amnesia/. An assessment of the residual risks for re-identification of participants in pseudonymised data sets should be performed. Under the General Data Protection Regulation (GDPR) in Europe there is an obligation on the data controller to carry out a data protection impact assessment (DPIA), to 'evaluate the origin, nature,

		particularity, and severity' of the 'risk to the rights and freedoms of natural persons' before processing personal data. The impact assessment 'should include the measures, safeguards and mechanisms envisaged for mitigating' the identified risks ***Annex 3***: Example of clinical trial DPIA content Services to support pseudonymisation of data sets, which could range from simple guidance, through consultancy, and on to performing and documenting the de-identification process, should be established.
Data Preparation	Selection of appropriate data standards has an impact on the secondary use. For instance, meta-analyses require alignment of data standards (for instance CDISC for clinical trial data). This is critical to the success of IPD sharing. Besides the IPD sets, other clinical trial data objects (metadata) should be made available for sharing (e.g., protocols, clinical study reports, statistical analysis plans, blank consent forms) to allow a full understanding of any data set.	Data and coding standards should be built into any trial's data design prospectively, from the beginning of the trial. While many scientific fields have developed and adopted common data standards, others have not. In such cases, the data sharing plan may indicate that no consensus data standards exist for the scientific data and metadata to be generated, preserved, and shared. If no standards exist, then the data should be managed so that is convertible to Study Data Tabulation Model (SDTM) format ***Annex 4***: Health data standards Consider sharing trial documents (study protocol and reports, statistical analysis plan, blank consent form, and other related documents that could help the reanalysis or understanding of the data)
DEPOSIT, STORAGE AND ACCESS		
Data storage (repositories)	Search a suitable repository. Data repositories have the potential to play an important role in the	Provide the name of the repository where scientific data and metadata arising from the project will be stored.

	effective and safe sharing of clinical study data. This is because they can provide a stable, long-term, secure environment? for the data, improve the security and quality of archiving through active data curation, increase the discoverability of data through the application of metadata schemes.	Data sharing requirement may vary between funding bodies and Data Holding Organisations (DHO). Data deposition repositories for suitable datasets can be searched using online data repository finder https://edctpknowledgehub.tghn.org/data-sharing-toolkit/repository-finder/ An evaluation of repositories for sharing individual-participant data from clinical studies analysed, in 2019, the landscape of data repositories to create a detailed description of available repositories and assess their suitability for hosting data from clinical studies, from the perspective of the clinical researcher (14). However, the landscape has evolved as the EU Regulations (GDPR, EHDS) will require the development of European, GDPR- and EHDS-compliant repositories, both from the technical (secure processing environment) and from the procedural point of view (Health Data Access Body). ***Annex 5.*** Examples of repositories that enable storage, sharing, discoverability, re-use of the IPD and associated documents from clinical studies When selecting a repository to manage and share data, here are some desirable characteristics to look for: **Unique Persistent Identifiers**: Assigns datasets a citable, unique persistent identifier, such as a digital object identifier (DOI) **Long-Term Sustainability**: Has a plan for long-term management of data, including maintaining integrity, authenticity, and

		availability of datasets; building on a stable technical infrastructure **Metadata**: Ensures datasets are accompanied by metadata to enable discovery, reuse, and citation of datasets, using schema that are appropriate to, and ideally widely used across, the community(ies) the repository serves. **Quality Assurance**: Provides, or has a mechanism for others to provide, expert curation and quality assurance to improve the accuracy and integrity of datasets and metadata. **Easy Access**: Provides broad, equitable, and maximally open access to datasets and their metadata in a timely manner after submission, consistent with legal and ethical limits required to maintain privacy and confidentiality, and protection of other sensitive data. **Broad and Measured Reuse**: Makes datasets and their metadata available with broadest possible terms of reuse; and provides the ability to measure attribution, citation, and reuse of data (i.e., through assignment of adequate metadata and unique persistent identifiers **Clear Use Guidance**: Provides accompanying documentation describing terms of dataset access and use (e.g., particular licenses, need for approval by a data use committee or ethics committee, fees). **Security and Integrity**: Has documented measures in place to meet generally accepted criteria for preventing unauthorized access to, modification of, or

		release of data, with levels of security that are appropriate to the sensitivity of data. **Confidentiality**: Has documented capabilities for ensuring that administrative, technical, and physical safeguards are employed to comply with applicable confidentiality, risk management, and continuous monitoring requirements for sensitive data. **Common Format**: Allows datasets and metadata downloaded, accessed, or exported from the repository to be in widely used, preferably non-proprietary, formats consistent with those used in the community(ies) the repository serves. **Provenance**: Has mechanisms in place to record the origin, chain of custody, and any modifications to submitted datasets and metadata. **Retention Policy**: Provides documentation on policies for data retention within the repository.
Discoverability (Metadata)	A metadata dictionary should be developed to make the clinical trial datasets findable. Evaluate if sufficient metadata will be available to enable data reuse.	Metadata is “data about the data” which describe the properties of the data set. Metadata should be structured so that it can be machine-read (e.g. by an Internet search engine) to maximize findability.

Data Access (rights, types and management of access)	IPD should be made available as soon as reasonably possible Check and clearly mention under which terms and conditions the data will be shared. Make sure that there is a licence, and that the licence gives you permission to do what you intend to.	Host organisations (e.g. Institute of Higher Education) may be able to provide funds for routine data sharing activities, e.g. ongoing maintenance of a data sharing system Responsibilities of staff for data sharing should be determined and funding should be sourced

5. DATA SHARING PLAN CHECKLIST

This checklist provides a step-by-step guide to determining whether your data sharing plan aligns appropriately with the guidelines provided. It includes elements which you should consider thoroughly in order to confirm that your data sharing plan complies with the EU and international funders expectations (*Annex 6).*

6. CONCLUSIONS

Many funding bodies require the data sharing plan or statements about data sharing for secondary use by the grant applicants in order to make their data available. The implementation of the guidelines outlined in this document will foster compliance with funder policies and expectation on data sharing. Researchers using these guidelines are encouraged to share their experience to inform future updates of this guidance.

7. REFERENCES

1. Reichman JH. Rethinking the Role of Clinical Trial Data in International Intellectual Property Law: The Case for a Public Goods Approach. SSRN Journal [Internet]. 2009 [cited 2024 Aug 14]; Available from: http://www.ssrn.com/abstract=1433392

2. Bauchner H, Golub RM, Fontanarosa PB. Data Sharing: An Ethical and Scientific Imperative. JAMA. 2016 Mar 22;315(12):1238.

3. Nosek BA, Alter G, Banks GC, Borsboom D, Bowman SD, Breckler SJ, et al. SCIENTIFIC STANDARDS. Promoting an open research culture. Science. 2015 Jun 26;348(6242):1422–5.

4. Nosek BA, Alter G, Banks GC, Borsboom D, Bowman SD, Breckler SJ, et al. Promoting an open research culture. Science. 2015 Jun 26;348(6242):1422–5.

5. Sharing and reuse of health-related data for research purposes: WHO policy and implementation guidance [Internet]. [cited 2024 Aug 14]. Available from: https://www.who.int/publications/i/item/9789240044968

6. Kiley R, Peatfield T, Hansen J, Reddington F. Data Sharing from Clinical Trials — A Research Funder's Perspective. N Engl J Med. 2017 Nov 16;377(20):1990–2.

7. ClinicalStudyDataRequest.com [Internet]. [cited 2024 Aug 14]. Available from: https://www.clinicalstudydatarequest.com/

8. Krumholz HM, Waldstreicher J. The Yale Open Data Access (YODA) Project — A Mechanism for Data Sharing. New England Journal of Medicine. 2016 Aug 4;375(5):403–5.

9. Bierer BE, Li R, Barnes M, Sim I. A Global, Neutral Platform for Sharing Trial Data. New England Journal of Medicine. 2016 Jun 23;374(25):2411–3.

10. GO FAIR [Internet]. [cited 2024 Aug 14]. FAIR Principles. Available from: https://www.go-fair.org/fair-principles/

11. 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 Editors. PLoS Med. 2017 Jun 5;14(6):e1002315.

12. 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 Dec;18(1):319.

13. Smith CT, Hopkins C, Sydes M, Woolfall K, Clarke M, Murray G, et al. Good practice principles for sharing individual participant data from publicly funded clinical trials. Trials. 2015 Dec;16(S2):O1.

14. Banzi R, Canham S, Kuchinke W, Krleza-Jeric K, Demotes-Mainard J, Ohmann C. Evaluation of repositories for sharing individual-participant data from clinical studies. Trials. 2019 Dec;20(1):169.

15. Kiley R, Peatfield T, Hansen J, Reddington F. Data Sharing from Clinical Trials - A Research Funder's Perspective. N Engl J Med. 2017 Nov 16;377(20):1990–2.

16. 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 Dec;7(12):e018647.

17. Kuntz RE, Antman EM, Califf RM, Ingelfinger JR, Krumholz HM, Ommaya A, et al. Individual Patient-Level Data Sharing for Continuous Learning: A Strategy for Trial Data Sharing. NAM Perspect. 2019;2019.

18. VACCELERATE [Internet]. [cited 2024 Apr 18]. Available from: https://vaccelerate.eu/

19. Staunton C, Slokenberga S, Mascalzoni D. The GDPR and the research exemption: considerations on the necessary safeguards for research biobanks. Eur J Hum Genet. 2019 Aug;27(8):1159–67.

20. 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 Dec 14;7(12):e018647.

21. Chan AW, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, et al. SPIRIT 2013 Statement: Defining Standard Protocol Items for Clinical Trials. Ann Intern Med. 2013 Feb 5;158(3):200.

8. ANNEX 1: EXAMPLE OF DESCRIPTION OF DATA SHARING WITHIN TRIAL PROTOCOLS

"In line with the EU data sharing policy, individual subject-level data will be shared with the scientific community (either as anonymised or pseudonymised data sets), while maintaining the integrity and privacy of the trial participants and in compliance with the EU General Data Protection Regulation (GDPR) and national or local rules. Data and other trial documents should be made available through an appropriate data repository, helping to ensure that the data objects are properly prepared, are available in the longer term, are stored securely and subject to rigorous governance. The terms and conditions of data transfers to a repository and the data sharing process shall be subject to specific data processing agreements to be established between the concerned parties as well as to a specific data sharing plan, where the details are specified".

9. ANNEX 2: CONSIDERATIONS FOR CONSENT ON DATA SHARING FOR FURTHER RESEARCH PURPOSES

I. Request for participation in future research

II. What is future research?

III. What is “pseudonymised” (OPTION 1)/ “anonymized” data (OPTION 2)

IV. What are the benefits of sharing my data?

V. Which categories of data are involved? (clinical data, laboratory records, genetic data)

VI. How will my data be shared?

- Anonymised/pseudonymized
- Within the European Economic Area (EEA) or outside
- Location of the data repository
- Time of storage
- Conditions to access

VII. What are the risks and how will my privacy be protected?

VIII. What if I want to withdraw from future research?

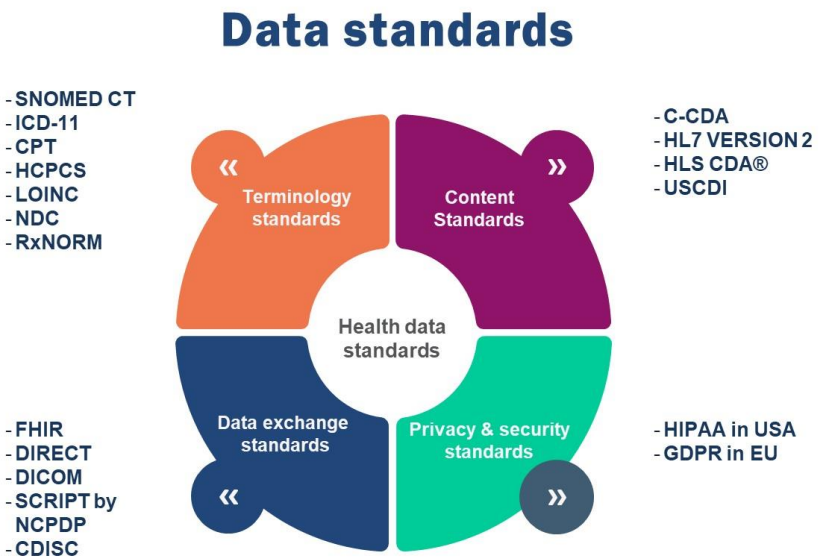
IX. How will I be informed about the results?

X. What are my rights and how can I exercise my rights?

10. ANNEX 3: EXAMPLE OF CLINICAL TRIAL DPIA CONTENT

1. Basic trial information and setting
2. Roles of Data Controller and Data Processors
3. Biobanking specifics
4. Separate research project on “Further Research”
5. Data of staff at clinical trial site and at Sponsor’s office
6. Collection of data
7. Use of data
8. Storage of data
9. Details of data deletion
10. Characterization of participants’ personal data
11. Sharing of participants’ personal data
12. Scope of data processing
13. Purpose of data processing steps
14. Security measures in place
15. International Transfer Impact Assessment (ITIA)
16. Description of EDC system
17. Legal basis
18. Data exit strategy
19. Consultation of external partners
20. Assessment of necessity and proportionality
21. Rights of participants
22. Identification and assessment of risks
23. Signatures of the Data Protection Officer and the Principal Coordinating Investigator

11. ANNEX 4: KEY HEALTH DATA STANDARDS

<table>
<tr><th colspan="2">Key health data Standards</th></tr>
<tr><td>Terminology Standards
Terminology standards utilize unique codes and classification systems to represent a wide range of health concepts and purposes</td><td rowspan="4"> 
</td></tr>
<tr><td>Content Standards
Content standards focus on defining the structure and data types of electronic medical documents. These standards ensure that medical data is properly organized and represented in a clear and easily understandable form</td></tr>
<tr><td>Data Exchange Standards
Data exchange standards establish the information flow between health systems. These standards ensure interoperability by specifying formats, document architecture, methods, APIs, and other necessary components for effective data exchange</td></tr>
<tr><td>Privacy and Security Standards
Privacy and security standards in healthcare are of utmost importance to protect sensitive health information from unauthorized access and ensure compliance with regulations such as HIPAA (Health Insurance Portability and Accountability Act) and EU-GDPR</td></tr>
</table>

12. ANNEX 5 REPOSITORIES FOR DATA SHARING IN CLINICAL RESEARCH (some examples)

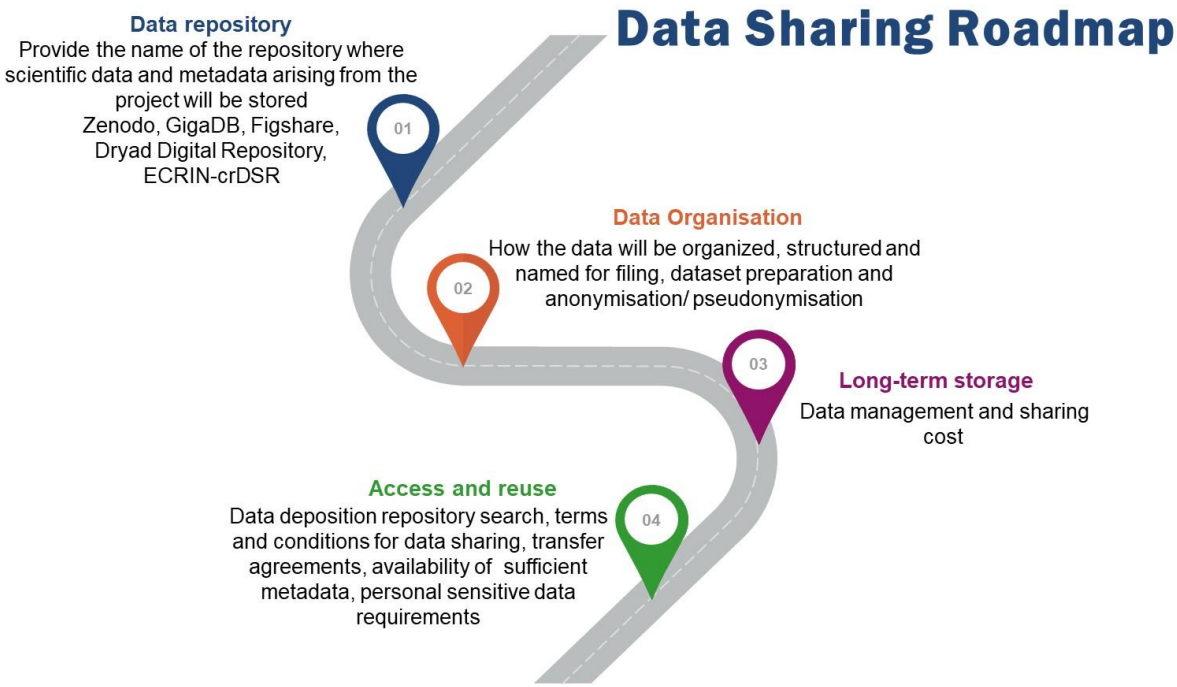
In 2019, Banzi et al., (14) made an assessment of the suitablity of different repositories for hosting clinical trial data.

	Guidelines for upload and storage	De-identification	Data quality control	Contract for upload and storage	Application of metadata	Application of identifiers	Flexibility of Access	Long term Preservation
Dryad	■	■	■	■	■	■	▒	■
Swedish National Data Service	■	◪	■	■	■	■	■	■
Drum	■	■	■	■	■	■	▒	▒
EASY	▒	■	▒	■	■	■	■	■
FigShare	▒	■	■	□	■	■	■	▒
ICPSR	■	□	■	■	■	■	■	▒
NDCT NIMH	■	■	■	■	■	■	□	◪
NDACAN (Child Abuse)	■	■	■	▒	■	■	□	◪
NIH BioLINCC	■	■	■	▒	■	◪	□	◪
Edinburgh DataShare	■	◪	▒	▒	■	■	■	■
Vivli	■	■	◪	■	◪	◪	▒	◪
B2Share	▒	□	□	◪	■	■	▒	▒
Open Science Framework	◪	◪	◪	◪	◪	■	■	▒
Project Datasphere	■	■	◪	■	□	□	□	◪
Zenodo	□	□	□	□	▒	■	■	■
NIDDK	■	■	▒	■	□	□	□	◪
ITN Trialshare	◪	■	◪	◪	▒	◪	□	◪
CancerData.Org	◪	◪	□	◪	■	■	□	◪
WWARN	■	▒	■	▒	□	□	▒	◪
Melanoma MMP	▒	◪	■	◪	◪	◪	□	◪
ProAct	■	■	▒	◪	▒	◪	□	◪
FreeBird	▒	▒	◪	◪	□	□	□	◪
EBCTCG	□	◪	■	◪	▒	□	□	◪
UMIN	▒	□	□	◪	◪	◪	□	◪
TBI-IMPACT	□	□	□	◪	◪	□	□	□

Legend

■ Demonstrated
□ Not demonstrated
▒ Partially demonstrated
◪ Missing or partial information available

13. ANNEX 6: DATA SHARING ROADMAP



14. ANNEX 7: DATA SHARING PLAN CHECKLIST