Training the Next Generation of Researchers in Clinical Data Sharing

Ulrich Mansmann
IBE, LMU
ulrich.mansmann@lmu.de
OSIM webinar, 12th of Dezember 2025

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

Overview

Motivation: Why does data sharing matter

Forms of data sharing: Claims data, Clinical routine data, registries, cohorts, clinical trials

History of clinical trial data sharing

SHARE-CTD: Building elements

How to share data: Practical implications

Challenges: Systemic and individual, data sharing stories

Careers: Risks and opportunities

Future and vision

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

Why (clinical trial) data sharing matters in medical research?

- Epistemic reason: Shared data reduces bias, allows replication, enables meta-analysis (IPD-MA). Openness to checking and re-analysis is a scientific virtue. (Karl Popper)
- Ethical reason: Patients / persons contribute data not just for one study, but in hope of broader societal benefit.
- Efficiency reason: Prevents duplication of effort, accelerates discovery, fosters innovation. There should be an economic intuition why sharing is efficient (Amartya Kumar Sen)
- Facts become black boxes when their making is not visible and open to inspection (Bruno Latour)
- Communality and organized skepticism: Two of Robert Merton's norms of science.
- Limits: Patient privacy, legal frameworks (GDPR, HIPAA), intellectual property, career incentives.
- Public reasoning: Data sharing expands the capabilities of the public to engage in informed debate, thereby deepening democracy.

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

Settings of data sharing in medicine

- Claims data are electronic records of healthcare services providers submit to insurers for payment
- Clinical routine data is health information collected automatically during everyday patient care (like GP visits or hospital stays) as part of standard operations
- Registries are organized databases collecting standardized info about groups with specific diseases, conditions, treatments, or exposures
- Cohorts are groups of people sharing a common trait (like birth year, exposure, or disease) tracked over time in longitudinal studies to understand how factors (diet, environment, genetics) affect long-term health outcomes, disease development, or treatment effectiveness.
- Clinical trials data is the comprehensive collection of information, results, and documents from controlled studies testing new drugs, devices, or treatments in humans.

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

History of clinical trial data sharing

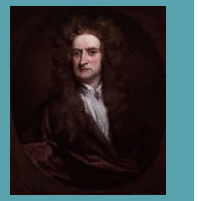
- Early Foundations (1970s-1990s): Openness as an ideal, not a practice
- First structural changes (200-2010): Registration and results transparency (CONSORT)
- The push for data sharing (2010-2014): Policy pressure and Ethical framing
- Institutionalisation (2015-2018): Platforms, policies, and legal constraints
- Regulatory complexity (GDPR, 2018): new constraints on data reuse, influencing cross-border sharing and increasing the need for governance frameworks, controlled access, and data de-identification standards.
- Recent period (2019 – present): Increasing maturity, uneven practice
- Emerging developments: Synthetic data and privacy-preserving analytics (federated learning, secure multiparty computation); Calls for harmonized global governance, rather than fragmented national or platform-specific rules; Shift toward seeing clinical data sharing as part of public infrastructure, not a voluntary add-on.

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

“Standing on the shoulders of giants”



Isaac Newton, Wikipedia

In clinical research, data sharing is a modern manifestation of this principle:

- Accumulated knowledge: Not starting from scratch, test new hypotheses, validate findings, meta-analyses.
- Transparency and reproducibility: Scrutinize, confirm, challenge published results, prevent errors, ensure robust scientific advancement.
- Innovation: Explore secondary questions, develop prediction models, identify treatment effect heterogeneity.
- Efficiency: “Investment” of patient participation benefits with each new scientific inquiry—maximizing value and reducing redundant trials.
- Collective progress: Advance knowledge faster and more reliably by standing on the shared “data shoulders” of prior studies.

Instead of standing on the shoulders of giants only through published summaries and interpretations, researchers can directly use the original *building blocks*—the patient-level data—provided by their predecessors.

Funded by the European Union

SHARE·CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

What kind of science do we want the next generation to practice?

Two extreme scenarios:

- Science as private (individual) achievement (individual datasets, guarded results).
- Science as collective endeavour (knowledge commons, reproducibility, shared responsibility, institutionalized sharing).

- Clinical (trial) data sits at the heart of this tension:
  It is intimate, sensitive, yet critical for generalizable knowledge.
  Data is often protected beyond institutional walls.

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

Social sculpture – 7000 Oaks (Joseph Beuys, 1982 documenta 7)

Joseph Beuys' concept of social sculpture shows how collective action shapes society.

In his project *7000 Oaks*, many individual contributions grew into a collective work that has a lasting impact.

Data sharing works in a similar way: Individual data points are shared, merged, and give rise to something new. In this way, compartmented knowledge becomes collective progress.

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

Social sculpture – 7000 Oaks (Joseph Beuys, 1982 documenta 7)

Fabian Püschel (C, D)

- Common process of creation
- Communication, cooperation, common action
- Participation
- Co-creation
- Openness
- Sustainability
- Tree + basalt stone
- Create by shared small contributions a new societal progress.

(E)

- The stones = “data points”
- The trees = “new knowledge that grows from data”
- The community = “Researchers who create something together”

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

... plans mainly detail efforts towards achieving widely supported science goals, such as data accessibility and reproducibility....

Scientists see the danger that the plan opens doors for more political inferences into science.

Nature, 29th of August,
doi:
https://doi.org/10.1038/d41586-025-02770-w
(F)

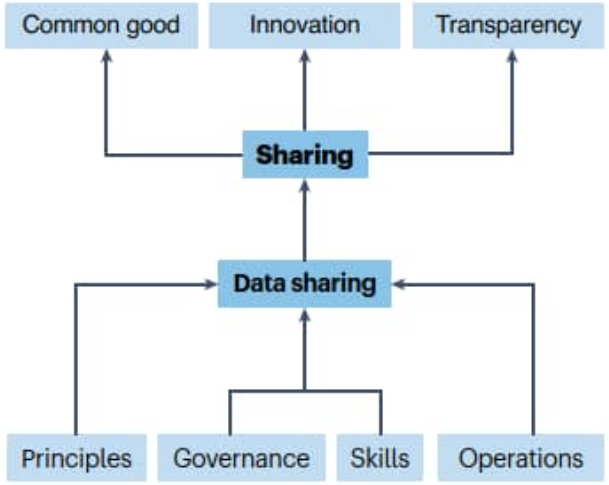
Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

Training the next generation

- Data sharing is a scientific virtue and a modern embodiment of falsifiability. It helps to consolidate and extend knowledge.
- Key competencies to train:
  - Technical literacy
  - Ethical sensitivity
  - Collaborative mindset
  - Scientific thinking
  - Critical skills
- Data sharing is not an “add-on”, but part of research integrity.



Nature Medicine, (Feb. 2023) V 29, 298-301, 299

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

SHARE-CDT

Comment

https://doi.org/10.1038/s41591-022-02080-y

Implementing clinical trial data sharing requires training a new generation of biomedical researchers

Ulrich Mansmann, Clara Locher, Fabian Prasser, Tracey Weissgerber, Ulrich Sax, Martin Posch, Evelyne Decullier, Ioana A. Cristea, Thomas P. A. Debray, Leonhard Held, David Moher, John P. A. Ioannidis, Joseph S. Ross, Christian Ohmann & Florian Naudet

Nature Medicine (Feb. 2023) V 29, 298-301, 299

SHARE-CTD: Start 01.01.2024, End: 31.12.2027

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

SHARE-CTD

How to provide data for sharing[1] – How to use shared data[2]

Good practice of Data Sharing – Impact of Data Sharing

The theory of CTDS (meta-science) – The practice of CTDS

Hierarchy of competences

CTDS: Clinical trial data sharing

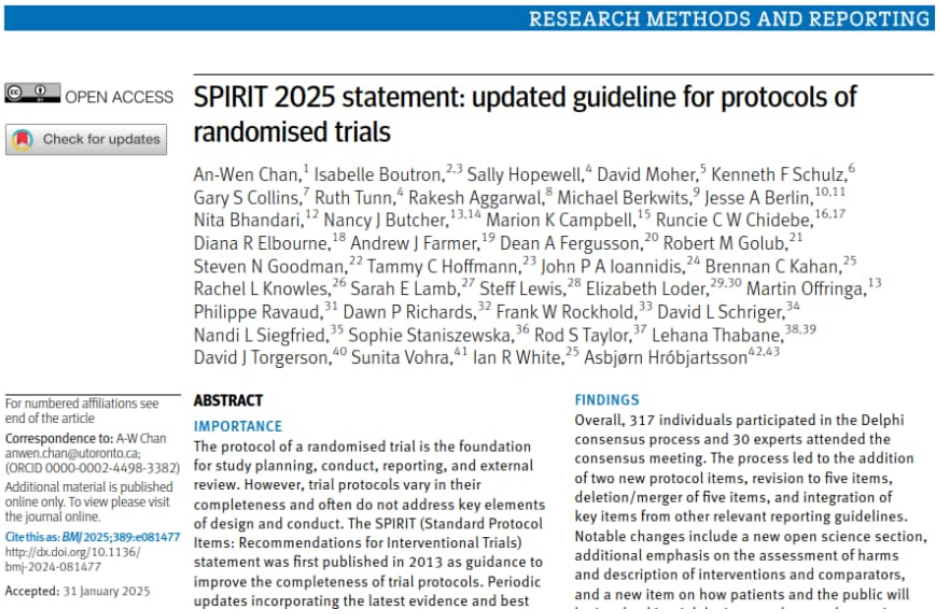
1: Tai KH et al (2025) Key Concepts in Clinical Epidemiology: FAIRification of Biomedical Research Data. J Clin Epidemiol. doi: 10.1016/j.jclinepi.2025.111920.
2. Varvara G et al (2025) Key Concepts in Clinical Epidemiology: Reusing clinical trial data to consolidate and advance medical knowledge, J Clin Epidemiol, doi: 10.1016/j.jclinepi.2025.111984.

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

Challenges: SPIRIT 2025 & CONSORT 2025



RESEARCH METHODS AND REPORTING

OPEN ACCESS

Check for updates

SPIRIT 2025 statement: updated guideline for protocols of randomised trials

An-Wen Chan,[1] Isabelle Boutron,[2,3] Sally Hopewell,[4] David Moher,[5] Kenneth F Schulz,[6] Gary S Collins,[7] Ruth Tunn,[4] Rakesh Aggarwal,[8] Michael Berkwits,[9] Jesse A Berlin,[10,11] Nita Bhandari,[12] Nancy J Butcher,[13,14] Marion K Campbell,[15] Runcie C W Chidebe,[16,17] Diana R Elbourne,[18] Andrew J Farmer,[19] Dean A Fergusson,[20] Robert M Golub,[21] Steven N Goodman,[22] Tammy C Hoffmann,[23] John P A Ioannidis,[24] Brennan C Kahan,[25] Rachel L Knowles,[26] Sarah E Lamb,[27] Steff Lewis,[28] Elizabeth Loder,[29,30] Martin Offringa,[13] Philippe Ravaud,[31] Dawn P Richards,[32] Frank W Rockhold,[33] David L Schriger,[34] Nandi L Siegfried,[35] Sophie Staniszewska,[36] Rod S Taylor,[37] Lehana Thabane,[38,39] David J Torgerson,[40] Sunita Vohra,[41] Ian R White,[25] Asbjørn Hróbjartsson[42,43]

For numbered affiliations see end of the article

Correspondence to: A-W Chan anwen.chan@utoronto.ca; (ORCID 0000-0002-4498-3382)

Additional material is published online only. To view please visit the journal online.

Cite this as: *BMJ* 2025;389:e081477 http://dx.doi.org/10.1136/bmj-2024-081477

Accepted: 31 January 2025

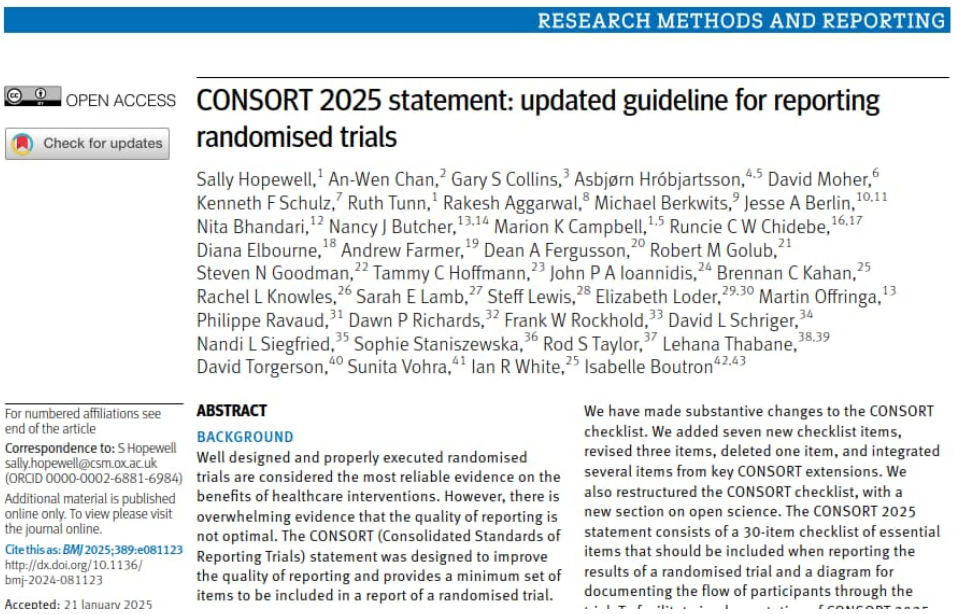
ABSTRACT

IMPORTANCE

The protocol of a randomised trial is the foundation for study planning, conduct, reporting, and external review. However, trial protocols vary in their completeness and often do not address key elements of design and conduct. The SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) statement was first published in 2013 as guidance to improve the completeness of trial protocols. Periodic updates incorporating the latest evidence and best

FINDINGS

Overall, 317 individuals participated in the Delphi consensus process and 30 experts attended the consensus meeting. The process led to the addition of two new protocol items, revision to five items, deletion/merger of five items, and integration of key items from other relevant reporting guidelines. Notable changes include a new open science section, additional emphasis on the assessment of harms and description of interventions and comparators, and a new item on how patients and the public will



RESEARCH METHODS AND REPORTING

OPEN ACCESS

Check for updates

CONSORT 2025 statement: updated guideline for reporting randomised trials

Sally Hopewell,[1] An-Wen Chan,[2] Gary S Collins,[3] Asbjørn Hróbjartsson,[4,5] David Moher,[6] Kenneth F Schulz,[7] Ruth Tunn,[1] Rakesh Aggarwal,[8] Michael Berkwits,[9] Jesse A Berlin,[10,11] Nita Bhandari,[12] Nancy J Butcher,[13,14] Marion K Campbell,[1,5] Runcie C W Chidebe,[16,17] Diana Elbourne,[18] Andrew Farmer,[19] Dean A Fergusson,[20] Robert M Golub,[21] Steven N Goodman,[22] Tammy C Hoffmann,[23] John P A Ioannidis,[24] Brennan C Kahan,[25] Rachel L Knowles,[26] Sarah E Lamb,[27] Steff Lewis,[28] Elizabeth Loder,[29,30] Martin Offringa,[13] Philippe Ravaud,[31] Dawn P Richards,[32] Frank W Rockhold,[33] David L Schriger,[34] Nandi L Siegfried,[35] Sophie Staniszewska,[36] Rod S Taylor,[37] Lehana Thabane,[38,39] David Torgerson,[40] Sunita Vohra,[41] Ian R White,[25] Isabelle Boutron[42,43]

For numbered affiliations see end of the article

Correspondence to: S Hopewell sally.hopewell@csm.ox.ac.uk (ORCID 0000-0002-6881-6984)

Additional material is published online only. To view please visit the journal online.

Cite this as: *BMJ* 2025;389:e081123 http://dx.doi.org/10.1136/bmj-2024-081123

Accepted: 21 January 2025

ABSTRACT

BACKGROUND

Well designed and properly executed randomised trials are considered the most reliable evidence on the benefits of healthcare interventions. However, there is overwhelming evidence that the quality of reporting is not optimal. The CONSORT (Consolidated Standards of Reporting Trials) statement was designed to improve the quality of reporting and provides a minimum set of items to be included in a report of a randomised trial.

We have made substantive changes to the CONSORT checklist. We added seven new checklist items, revised three items, deleted one item, and integrated several items from key CONSORT extensions. We also restructured the CONSORT checklist, with a new section on open science. The CONSORT 2025 statement consists of a 30-item checklist of essential items that should be included when reporting the results of a randomised trial and a diagram for documenting the flow of participants through the

Process transparency

Results transparency

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

How to provide data for sharing?
SPIRIT 2025 & CONSORT 2025

SPIRIT 2025 (item 4-6)	CONSORT 2025 (item 4)
In the protocol phase, researchers are now expected to pre-specify: • Whether they plan to share data, code, materials • How they will ensure accessibility and ethical reuse • Which repositories will be used (e.g., Dryad, Zenodo, Vivli) • Structured prospective declaration Why this matters: Pre-specifying open science commitments discourages post hoc decisions that could hide unfavorable data or methods.	In the trial report, researchers must state: • What was actually shared • Where it can be found (links to repositories, DOIs) • If sharing was not possible, a clear justification is required • Transparent reporting Why this matters: Providing clear structures and information enables optimal use of data for follow-up research.

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

How to provide data for sharing?

The step *before* the actual sharing is crucial to ensure that data is truly reusable, understandable, and legally secure. The most important aspects of preparing data for data sharing are:

- Data quality and consistency
- Documentation (metadata)
- Anonymization and data protection
- Technical preparation
- Legal and ethical aspects
- Long-term accessibility

Overall, the FAIR principles should apply: Data should be findable, accessible, interoperable, and reusable.

Ka-Hin Tai et al. (2025) https://pubmed.ncbi.nlm.nih.gov/40774362/

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

Data Sharing Statement - DSS

- Availability: The dataset[s] generated and/or analyzed during the current study are [available from the corresponding author on reasonable request / publicly available at NAME OF REPOSITORY with DOI / not publicly available due to [reasons]].

- Conditions: Data will be shared [in anonymized form / under a data sharing agreement / upon ethical approval] and are intended for [research purposes only / academic use only].

- Who to ask: Data Requests should be directed to: [email address or contact person].

- Access: Access will be granted [upon approval by the principal investigator / within [X] weeks], subject to compliance with applicable data protection laws.

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

How to access and use data that is shareable?

The data request process should balance feasibility with openness.

Handling data formats, managing deployment processes, implementing secondary studies

Define your motivating for clinical trail data sharing in the DSA to avoid cherry picking if data is in your hand:

- Reproducing results, checking their robustness, post-publication review
- Clinical trial planning
- Analyzing clinical trial aspects: How does a center develop during a trial, etc.
- IPD meta –analyses
- (Biostatistical) methods development
- Validation activities (validating prediction models, …)

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

Data Sharing Agreement (DSA)

Understand the obligatory (generic) parts of a DSA are:

1. Purpose of Data Sharing → Cherry-picking and preregistration
2. Description of the Data
3. Legal and Ethical Compliance
4. Data Protection and Security
5. Intellectual Property and Ownership
6. Publication and Reporting
7. Duration and Termination
8. Liability and Indemnity
9. Governing Law and Dispute Resolution

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

How to use data that can be shared?

The steps *after* the getting access to the data…

Understanding conditions of use

Ensuring data security

Respecting participant privacy

Scientific integrity

Ethical responsibility

Planning dissemination

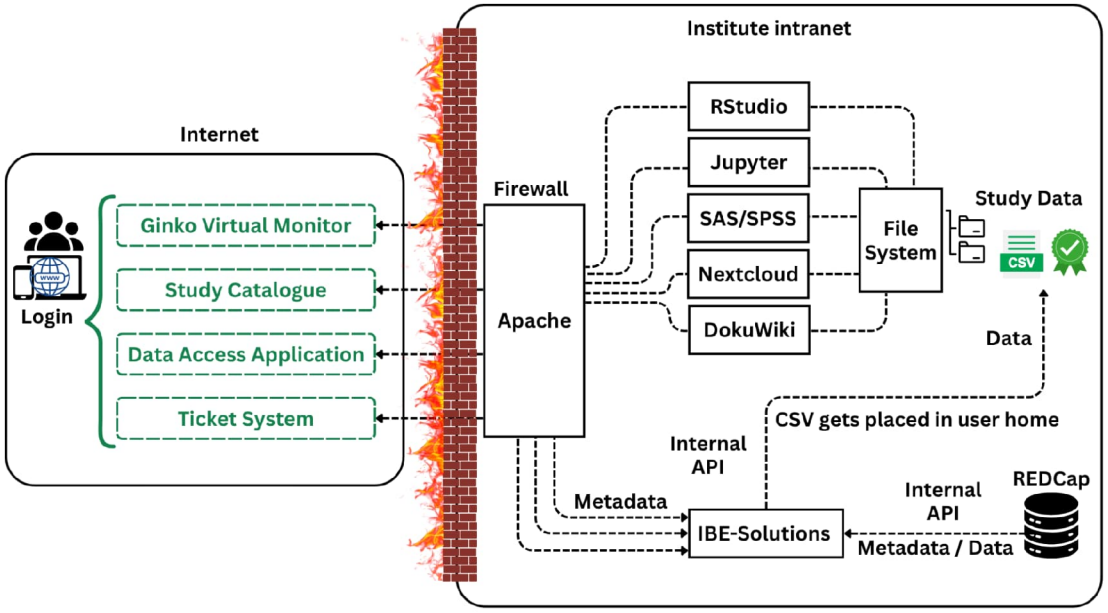
Think about compliance, security, ethics, transparency, and reciprocity—treat shared data not simply as a resource.

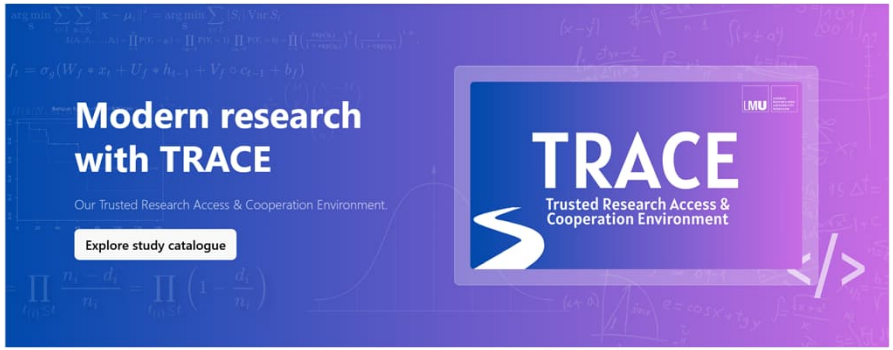
Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

Cheap local IT sharing infrastructure





Providing a technical infrastructure for CTDS
Providing a training platform for *how to provide data for sharing* as well as *how to use shared data.*

mmueller@ibe.med.uni-muenchen.de

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

Data Repositories: How to use data that can be shared?

CSDR, YODA, VIVLI, TRACE….

Principle	Guidance for repositories
Transparency	To be transparent about specific repository services and data holdings that are verifiable by publicly accessible evidence.
Responsibility	To be responsible for ensuring the authenticity and integrity of data holdings and for the reliability and persistence of its service.
User Focus	To ensure that the data management norms and expectations of target user communities are met.
Sustainability	To sustain services and preserve data holdings for the long-term.
Technology	To provide infrastructure and capabilities to support secure, persistent, and reliable services.

The TRUST Principles for data repositories

Lin et al. Scientific Data 7, 144 (2020)

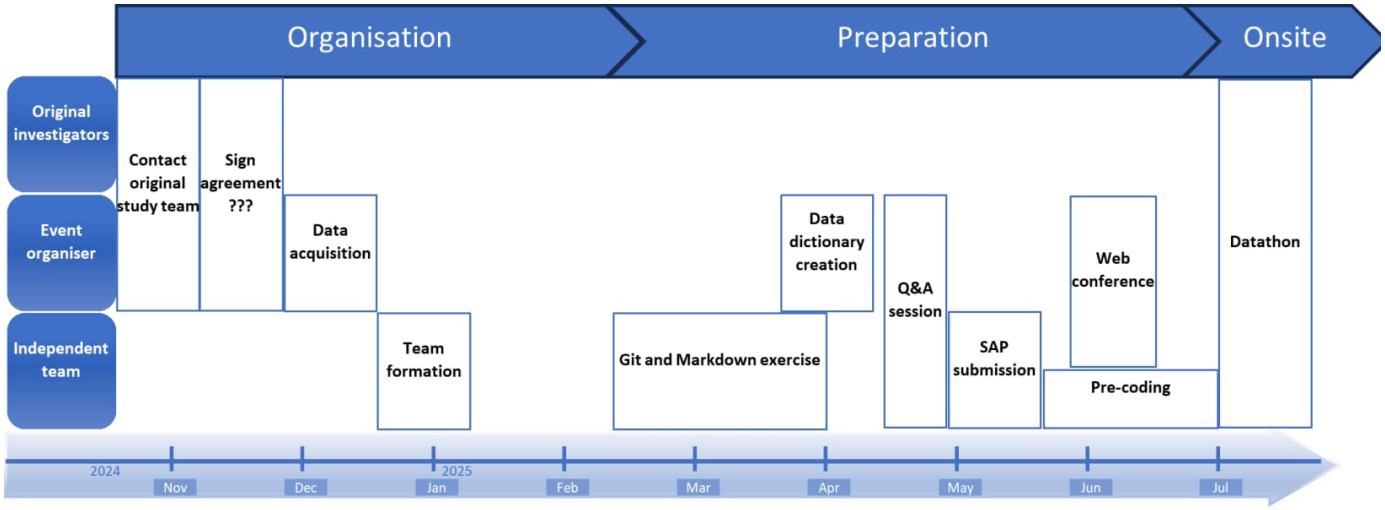
Academic and commercial communities are developing best practices for organizing and performing data sharing.

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

Datathons



(G)

Oxaliplatin added to fluorouracil-based preoperative chemoradiotherapy and postoperative chemotherapy of locally advanced rectal cancer (the German CAO/ARO/AIO-04 study): final results of the multicentre, open-label, randomised, phase 3 trial

Lancet Oncol 2015; 16: 979–89

Ka-Hin Tai, et al (2025)

Paper submitted to BMC Trials

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

Technics to be learned and trained

Transparent reanalysis of clinical trial data requires secure data handling, reproducible workflows, rigorous statistics, open sharing, and clear collaborative communication.

Data Management & Preparation	Transparency & Open Science Practices
Data access & governance; Data import & integration Data cleaning & preprocessing; Metadata management	Pre-registration & protocols; Sharing code & outputs Reporting standards; Reproducibility checks
Reproducible & Transparent Workflows	Communication & Collaboration
Version control; Pipeline automation; Containerization & environments; Documentation standards	Clear documentation; Interdisciplinary collaboration Visualization & interpretation
Statistical & Computational Competence	
Statistical methods; Simulation & sensitivity analysis; Computational efficiency	

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

IPD Meta-analysis

An individual patient data (IPD) meta-analysis is a type of systematic review in which the *raw, participant-level data* from each included study are collected, checked, and re-analyzed centrally, rather than relying only on published summary statistics (e.g., hazard ratios, means, odds ratios).

Feature	Classical Meta-analysis	IPD Meta-analysis
Data used	Published summaries	Raw participant data
Flexibility	Limited	Very high
Subgroup analyses	Weak / often unreliable	Strong and reliable
Data quality check	Not possible	Full checking possible
Bias control	Moderate	Best available
Resource needs	Low	High

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

Post publication review

Post-publication review (PPR) is highly relevant in clinical sciences as it provides an ongoing, transparent, and community-wide assessment of published research, complementing traditional pre-publication peer review to enhance the integrity, accuracy, and reliability of medical literature and clinical practice.

In essence, PPR serves as a vital, dynamic feedback mechanism that helps the clinical sciences community refine its body of knowledge and ensure that clinical practice is based on the most robust and reliable evidence possible.

Despite the existence of platforms like PubPeer and ResearchGate that facilitate PPR, systemic and cultural barriers imply that the process has not been widely or consistently adopted in the clinical sciences: Lack of incentives and recognition, time and resource constraints, high volume of literature, cultural and social factors, anonymity concerns, journal conflicts of interest, clinical implications, variability and inconsistency.

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

Incentives for Data Sharing

Scientific gains: Increased citations, collaborations, secondary analyses

Reputational benefits: greater transparency and credibility

Compliance with funder, regulator, or journal requirements. Data sharing can also accelerate Innovation and patient benefit, reinforcing the social value of research.

However, strong incentives require recognition of data sharing as a scholarly contribution, support for infrastructure, and safeguards to protect privacy and researchers' intellectual investment.

The core idea is to embed open science in the reward system—so that researchers gain tangible academic and professional recognition, not just moral credit, for sharing and reusing trial data.

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

Ideal narrative and current academic reward system

While the ideal narrative around CTDS stresses transparency, reproducibility, and societal benefit, the current academic reward system often lags behind:

Publication bias – High-impact journals tend to prioritize *novel* findings over replication studies or secondary analyses from shared data, which are often seen as less original.

Recognition gap – Data curators and secondary analysts frequently receive limited academic credit compared to primary trial investigators.

Career incentives – Hiring, promotion, and funding decisions are still driven by high-profile publications rather than contributions to open science.

Editorial caution – Journals may perceive re-analyses or secondary findings as carrying more methodological or interpretive risks.

So while data sharing is encouraged in principle, the reward structures in practice still discourage researchers who depend on high-impact outlets.

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

CTDS - Career Perspectives

Risks / Concerns	Opportunities / Benefits
Fear of being scooped – others may publish faster with “your” data – research parasites.	Visibility & citations – datasets get DOIs and can be cited; your work lives longer than a single paper.
Extra workload – preparing anonymized, well-curated datasets is time-consuming.	Skill development – data curation, harmonization, reproducibility → highly valued skills in academia, pharma,
Credit systems lag – tenure & hiring still emphasize papers, not shared datasets.	Alternative outputs – data papers (*Scientific Data*, *GigaScience*), reproducible workflows, software packages.
Unclear incentives – institutions may not reward openness.	Collaboration gateway – joining data consortia, IPD meta-analyses, international projects → more co-authorships.
Ethical/legal complexity – consent, GDPR/HIPAA, governance add uncertainty.	Ethical leadership – expertise in patient privacy & FAIR principles builds trust and opens leadership roles.
Loss of control – once data are public, usage is unpredictable.	Scientific stewardship – being known as a *trusted steward* enhances credibility and employability.

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

CTDS – our optimistic Career Perspectives

For the next generation, learning data sharing is not a burden but a passport. It is a passport to collaborations across borders, to credibility in the eyes of funders and journals, and to transferable skills that will outlast any single trial or dataset.

For early-career researchers, data sharing may look like a cost. But in reality, it is an investment — in skills, visibility, and credibility that will pay off across an entire career.

Practical guidance as an output of SHARE-CTD – added to some already existing

› BMC Med Res Methodol. 2016 Jul 8;16 Suppl 1(Suppl 1):73. doi: 10.1186/s12874-016-0171-x.

EFSPI/PSI working group on data sharing: accessing and working with pharmaceutical clinical trial patient level datasets – a primer for academic researchers

Rebecca Sudlow [1], Janice Branson [2], Tim Friede [3], David Morgan [4], Caroline Whately-Smith [5]

Affiliations + expand

PMID: 27410386 PMCID: PMC4943504 DOI: 10.1186/s12874-016-0171-x

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

Challenges and future directions

- Barriers: Systemic, individual, ethical, economical
- Incentive structures: career progression rewards publications, not data curation.
- Trust issues: fear of being scooped, misuse of data, loss of control.
- Training gap: very few PhD programs integrate “open science & data stewardship” as core modules.
- Possible remedies: Make CTDS to a norm (SPIRIT, CONSORT), embed data sharing principles into graduate curricula, reward datasets with citations (DOIs, credit systems).
- Teaching narrative: “Every dataset tells more than one story — share it so others can hear theirs.”

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

Vision

- Clinical data are gifts from patients.
- As researchers, we are not their owners but their guardians.
- The next generation must learn not only how to analyze data, but how to share them responsibly — because the future of medicine depends not just on what we discover, but on what we make discoverable.
- Closed science is efficient in a short run; open science is resilient in the long run.
- Connect science back to its social contract: trust, transparency, responsibility
- Lighttower: guidance, stewardship, and visibility.
  Illuminate the path for others by making data discoverable and usable.

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

Summary: The theory and the practice of CTDS

- Clinical trial data sharing is the practice of making individual participant data, study protocols, and analysis plans accessible to qualified researchers.
- In theory, it advances scientific transparency, reproducibility, and trust, enabling independent verification of results, meta-analyses, and new discoveries.
- In practice, data sharing requires robust governance frameworks to protect patient privacy, ensure ethical use, and balance openness with safeguards.
- While initiatives from regulators, journals, and funders promote sharing, challenges remain around standardization, consent, and incentives for researchers.
- Overall, data sharing is both a technical and cultural shift, aiming to improve the reliability and social value of clinical research.

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

The philosophical perspective

- Karl Popper's vision: science advances by critical scrutiny; data sharing is the modern embodiment of "falsifiability".
- Merton's norms of science: communality, universalism, disinterestedness, organized skepticism — data sharing operationalizes communality.
- Amarty Kumar Sen: Expand capabilities of the public
- Bruno Latour: Interconnection between science, society, and the environment (ANT: actor-network theory)
- Donna Haraway: Knowledge is situated. How is it produced? Motivates rich metadata and method transparency.

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

Many thanks

ulrich.mansmann@lmu.de

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

References for images

A http://www.derbymuseums.org/joseph-wright-gallery/

B from BMJ 331 (Dec 2005)

C https://daskunstbuch.at/2014/10/06/joseph-beuys-7000-eichen-der-versuch-einer-sozialokologischen-gemeinschaftsarbeit/

D Fabian Püschel, 7000 Eichen (Film), 2010 (15.10.2013), URL: http://vimeo.com/15749968

E https://commons.wikimedia.org/wiki/File:7000_Eichen-Stelen_05_2016-03-19.jpg

F Credit: Samuel Corum/Sipa/Bloomberg/Getty, https://www.nature.com/articles/d41586-025-02770-w

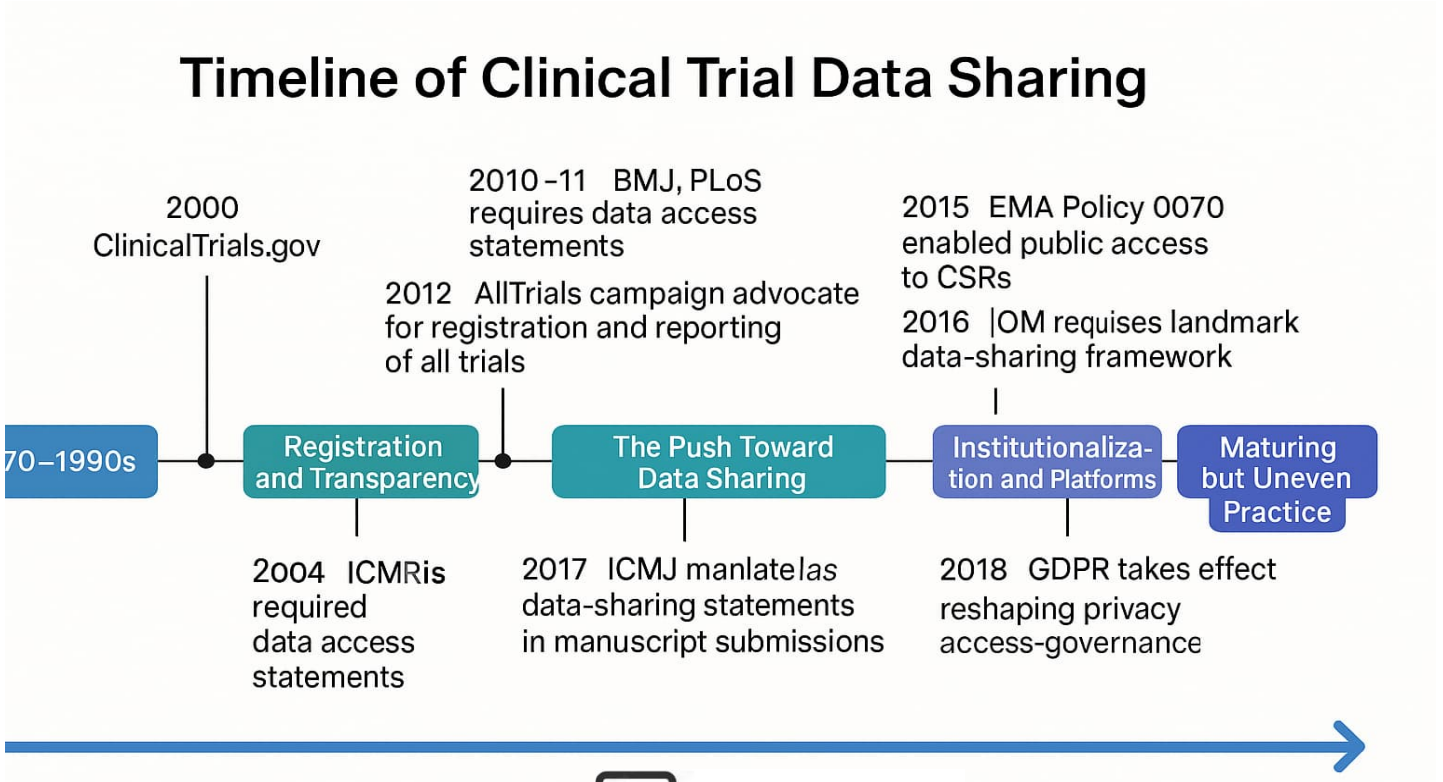
G https://share-ctd.eu/user/pages/01.home/D1_Bild1.jpg

Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

Short history of clinical trial data sharing

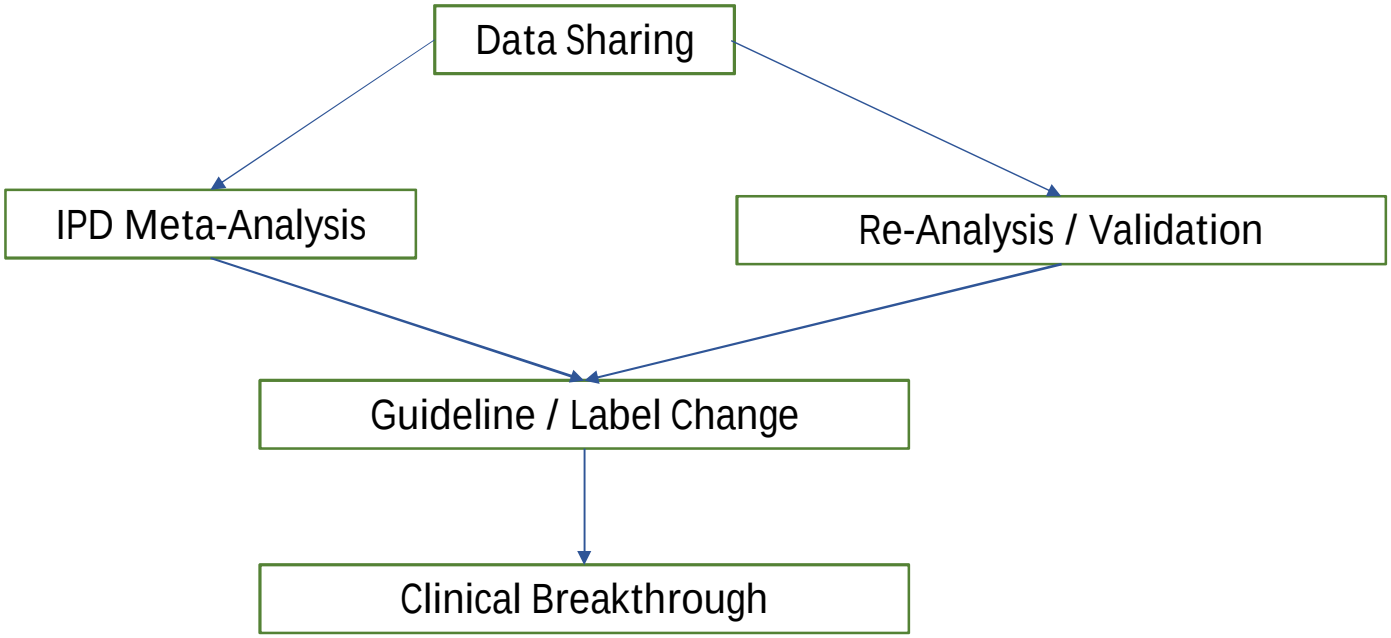


Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360

How I would like to use shared data and gain impact.



Funded by the European Union

SHARE · CTD
Cooperate to share and gain

Horizon-MSCA-2022-DN-01
Project: 101120360