



SHARE • CTD

Cooperate to share and gain

Clinical Data Sharing for Better Health Decisions

Rita Banzi

Mario Negri Institute for Pharmacological Research IRCCS

Milan, Italy



**Funded by
the European Union**

Horizon-MSCA-2022-DN-01

Project: 101120360

Declaration of interest

No personal financial or economic conflicts of interest related to this presentation.

My research is primarily supported by publicly funded Italian and European projects.

I do have intellectual conflicts of interest:

I am a fan of evidence-based health decisions, public and universal healthcare systems, and ethically and scientifically rigorous research.

Focus of the presentation

This will not be a technical presentation but rather an inspirational lecture on why clinical trial data sharing and research transparency matter for everyone involved in evidence-based decision-making, from researchers to policymakers and healthcare professionals.

Your PhD work is therefore highly valuable, as it contributes to improving the quality of health decisions.

11 years ago




The Institute of Medicine's 2015 report on transparency in clinical research

Strengths and limitations

Rita Banzi, Mario Negri Institute, Milan, Italy
rita.banzi@marionegri.it

21 May 2015, International Clinical Trial Day,
Trondheim, Norway



Sharing Clinical Trial Data

MAXIMIZING BENEFITS, MINIMIZING RISK

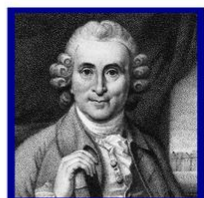
INSTITUTE OF MEDICINE
OF THE NATIONAL ACADEMIES

A couple of weeks ago

International Clinical Trials Day (ICTD) 2026

Stronger Together: Advanced Therapy Clinical Trials Without Borders

ICTD
International
Clinical Trials Day
20 May 2026



#ICTD2026

ecrin CZECRIN

Stronger Together: Advanced Therapy
Clinical Trials Without Borders

REGISTER NOW

20 May, 2026

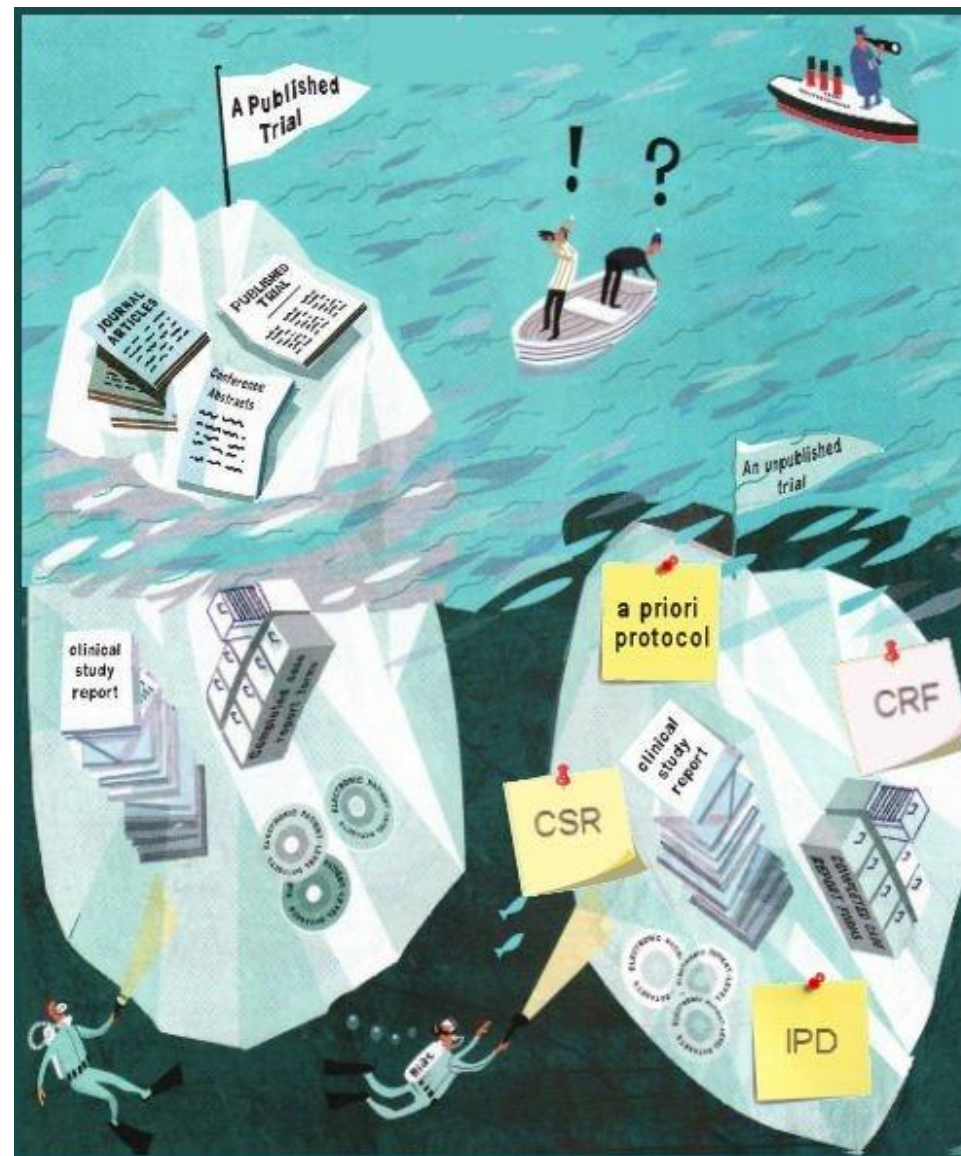
Empire Hall, Slovanský dům
Prague, Czech Republic

Advanced Therapy Medicinal Products (ATMPs)—such as gene therapies, cell therapies, and tissue-engineered products—pose challenges for

- Development
- Licensing
- Reimbursement

**Call for
radical transparency**

Health decision-makers are blind to the totality of evidence...



Doshi BMJ 2013

...and only see the bright side of the moon

Summary of findings 1. Summary of findings: publication and time to publication of clinical trials

Population: observational studies that followed a cohort of interventional clinical trials			
Settings: trial registries, ethics committees, funding organisations, protocol databases			
Intervention: publication			
Comparison: no publication			
Outcomes	Anticipated absolute and relative effects (95% CI, IQR, SD)	Number of studies	Comments
Publication¹	The prevalence of publication was 0.53 (95% CI 0.51 to 0.56).	203	Only just over half of all clinical trials included in this review have been published.
Time to publication			
Time from enrolment of first participant	The median (IQR) of the medians was 57.2 months (IQR 52.2).	10	The median time to publication from enrolment of the first participant was on average 4.8 years.
Time from completion of the trial ²	The median (IQR) of the medians was 25.6 months (IQR 11.6).	54	The median time to publication from completion of the trial was on average 2.1 years.
Effect of positive results on publication	OR 2.69 (95% CI 2.02 to 3.60)	19	Trials with positive results were more likely to be published than trials with negative or null results.
Effect of positive results on time to publication²	Adjusted HR 1.92 (95% CI 1.51 to 2.45)	4	Trials with positive results were published in less time than trials with negative or null results; on average, 2 and 2.6 years (median), respectively.
Effect of large sample size on publication	OR 1.92 (95% CI 1.33 to 2.77)	11	Trials with large sample sizes were more likely to be published than trials with small sample sizes.
Effect of large sample size on time to publication²	Adjusted HR 1.41 (95% CI 1.18 to 1.68)	7	Trials with large sample sizes were published in less time than trials with small sample sizes.
Effect of multicentre trials on publication	Adjusted OR 1.20 (95% CI 1.03 to 1.40)	2	Multicentre trials were more likely to be published than single-centre trials.
Effect of multicentre trials on time to publication²	Unadjusted HR 1.15 (95% CI 0.96 to 1.37)	3	There was no apparent difference in time to publication between multicentre and single-centre trials.
Effect of non-industry funding on publication	Adjusted OR 2.13 (95% CI 1.82 to 2.49)	14	Non-industry-funded trials were more likely to be published than industry-funded trials.
	Unadjusted OR 1.32 (95% CI 1.18 to 1.47)	83	
Effect of non-industry funding on time to publication²	Adjusted HR 1.46 (95% CI 1.15 to 1.86)	7	Non-industry-funded trials were published in less time than industry-funded trials.



**Cochrane
Library**

Cochrane Database of Systematic Reviews

Time to publication for results of clinical trials (Review)

Showell MG, Cole S, Clarke MJ, DeVito NJ, Farquhar C, Jordan V

Showell Cochrane Database of Systematic Reviews 2024
(update of Hopewell 2009)



SHARE · CTD
Cooperate to share and gain

Three examples

How a lack of transparency affects

- Patient safety
- Public funds
- Statistical validation

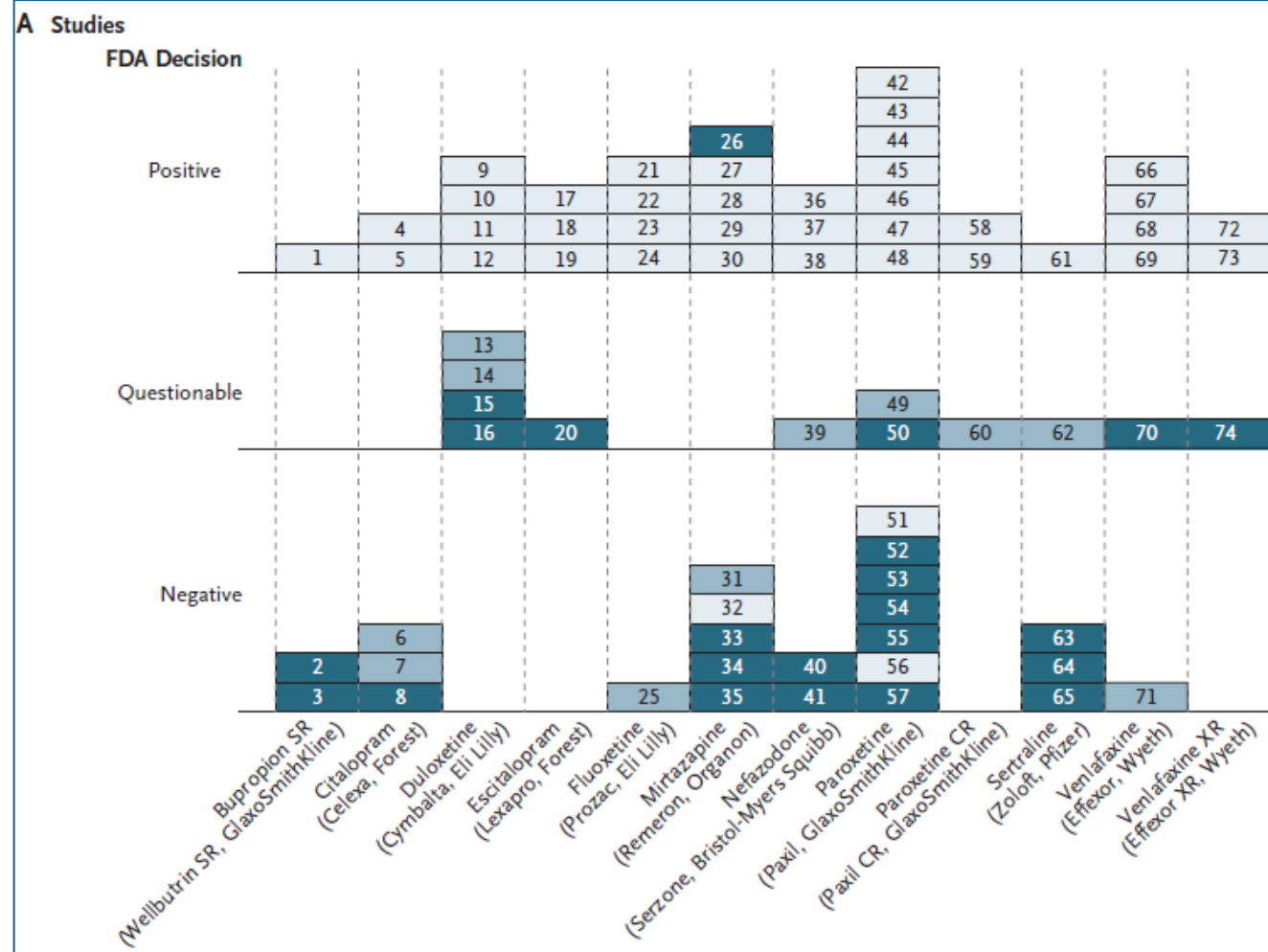
Example 1: antidepressants

SPECIAL ARTICLE

Selective Publication of Antidepressant Trials and Its Influence on Apparent Efficacy

Erick H. Turner, M.D., Annette M. Matthews, M.D., Eftihia Linardatos, B.S.,
Robert A. Tell, L.C.S.W., and Robert Rosenthal, Ph.D.

Published, agrees with FDA Published, conflicts with FDA Not published



Turner et al N Engl J Med 2008

Example 1: antidepressants

Risk–benefit ratio of antidepressants in adolescents

- overestimation of benefits
- underestimation of harms, including suicidal thoughts and behaviors.

When unpublished and selectively reported trial data were later considered, concerns emerged that the safety profile of these medications was less favorable than initially suggested.



Example 2: antivirals for influenza

Pharmaceuticals industry

This article is more than 12 years old

What the Tamiflu saga tells us about drug trials and big pharma

We now know the government's Tamiflu stockpile wouldn't have done us much good in the event of a flu epidemic. But the secrecy surrounding clinical trials means there's a lot we don't know about other medicines we take



**Ben Goldacre**
Thu 10 Apr 2014 08.00 CEST

 **Share**  **718**

 Prefer the Guardian on Google

Between 2006-07 and 2012-13, the UK Health Department spent £560 million on stockpiling two antiviral medicines for use in an influenza pandemic—£424 million on Tamiflu and £136 million on Relenza.

Example 2: antivirals for influenza

Tamiflu campaign – researchers, publishers, media

Open data campaign aimed to pressure companies into releasing the underlying clinical trial data for two globally stockpiled anti-influenza drugs

- WHO was recommending Tamiflu but had not vetted the underlying data.
- EMA approved Tamiflu, but had not vetted the underlying data.
- CDC was encouraging the use and stockpiling of Tamiflu on the basis of the 6-page manufacturer funded pooled analysis of 10 clinical trials, but had not vetted the underlying data (despite the FDA, which had vetted the underlying data, recognized the lack of efficacy on influenza complications).

Example 3: surrogate endpoints

Trialists and decision-makers like surrogate endpoints

- Depending on the disease or health area and definitions of a surrogate endpoint, between 17% and 78% of trials use surrogate endpoints as primary outcomes.
- Medicines are often licensed based on surrogate endpoints, especially in the case of accelerated and conditional approvals.
- Approximately 60% of new medicines and biologics approved by the US FDA in the last 2 decades are based on surrogate endpoints.

"When Is a Surrogate Endpoint a Good Surrogate?"

Example 3: validated surrogate endpoints

Validated surrogate endpoint should show utility as both

- **prognostic marker** (patient-level surrogacy) — *can the surrogate predict outcomes for individual patients?*
- **predictive marker** (trial-level surrogacy) — *can treatment effects on the surrogate predict treatment effects on the final endpoint?*

The optimal evidence base would include multiple RCTs with accessible participant-level data (IPD)

Level	Method	Pros	Cons
Levels 1 and 2	IPD bivariate meta-analysis using regression-based mixed models ^{29,30,36}	<ul style="list-style-type: none"> • Models allow for estimation of both the individual-level and trial-level association. • Naturally account for measurement error. • Less strong distributional assumptions at the within-trial level (compared with methods for aggregate data). • Can be used to adjust for effect modifiers available in IPD datasets. 	<ul style="list-style-type: none"> • Require IPD from all trials. • Assume bivariate normal distribution at the between-study level—sometimes a strong assumption as discussed below.

But this is rarely feasible, so regulatory and Health Technology Assessment agencies may endorse other approaches as **acceptable** solutions, but methodological recommendations vary considerably.

Are we doing enough to ensure clinical research data transparency?

WHO recommendations (2025)

Clinical Trials in Global Health 5

Reporting summary results in clinical trial registries: updated guidance from WHO

An-Wen Chan, Ghassan Karam, Justin Pymento, Lisa M Askie, Luiza R da Silva, Anna Laura Ross, Vasee Moorthy

The importance of publicly registering clinical trials and reporting substantial progress has been made with registering trials but remains uncommon despite expanding legislative and funding support and avoidable waste of resources, particularly for unpublished summary results in trial registries, reviews the current landscape for reporting results in registries. The 2025 WHO guidance includes researchers, patients, sponsors, funders, regulators, and funders. The guidance defines eight minimum items that are essential for all trials. Implementation of the WHO guidance by trial registries and endorsement by funders, regulators, legislators, research funders can help enhance the contribution of trials to scientific knowledge.

	Item name	Description	Components
1	Trial protocol	Most recent study protocol and full statistical analysis plan, including version number, date, and history of amendments	Document upload or digital object identifier
2	Completion status	When and why the trial ended or was stopped	Completion date and completion status (completed as expected or stopped early, with reason)
3	Dates of reporting results	Dates when results were reported in a journal or registry	Date and digital object identifier of first publication of primary results; date of submitting results to the registry; and date of posting of results in the registry
4	Participant flow	Progress of participants from study enrolment to primary analysis (ie, based on CONSORT flow diagram)	For each group, numbers of participants who were allocated to intervention, did not receive intervention as allocated (with reasons), were lost to follow-up, and were excluded from the main analysis of primary outcome (with reasons)
5	Participant characteristics	Characteristics of participants at baseline	For each group, descriptive summary statistics for age, sex or gender, and other relevant sociodemographic and study-specific characteristics
6	Outcome results	(a) Definition of each primary and secondary outcome	Specification (eg, primary outcome); measurement variable; analysis metric (eg, final value); method of aggregation (eg, mean); and timepoint of analysis
6	Outcome results	(b) Who is included in the analysis, and in which group, for each outcome	Definition of analysis population (eg, all randomly assigned participants in their original assigned group)
6	Outcome results	(c) Summary by group for each outcome and analysis population	Name of trial group; number of participants analysed; for binary data, the number of participants with the outcome event; and for continuous data, the average value (eg, mean or median), and measure of variability (eg, standard deviation or interquartile range)
6	Outcome results	(d) Comparison between groups for each outcome and analysis population	Names of groups being compared; effect measure (eg, odds ratio, relative risk, hazard ratio, risk difference, and difference in means); effect size; confidence interval and level (eg, 95% CI); and p value (optional)
7	Harms or adverse events	Unfavourable changes in health (eg, new or worsening symptom, abnormal laboratory finding) in each group, regardless of causal relation to the study intervention	Definition of harms analysis population (eg, all randomly assigned participants in their original assigned group); and numbers and percentages of participants by trial group for deaths, serious adverse events (overall and each type), and each type of adverse event by severity
8	Conflicts of interest	Financial and non-financial relationships that create conflicts of interest	Any conflicts of interest for trial steering committee members (describe or refer to journal publication)

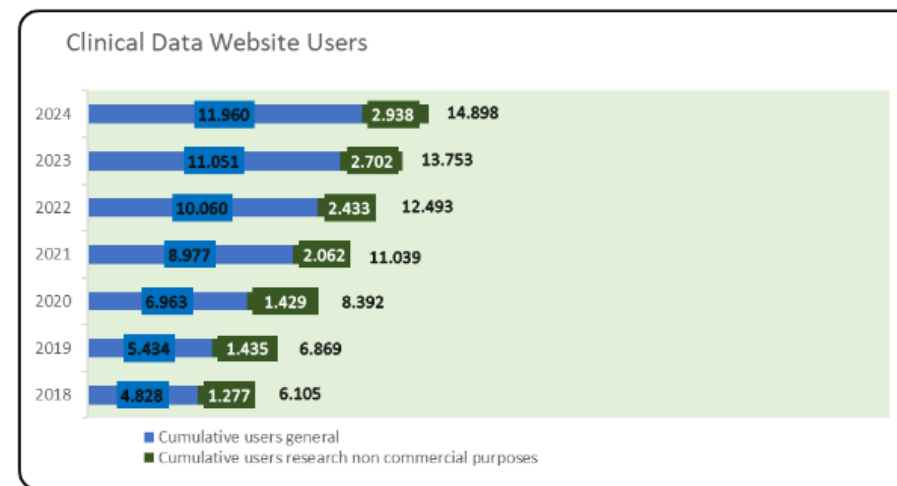
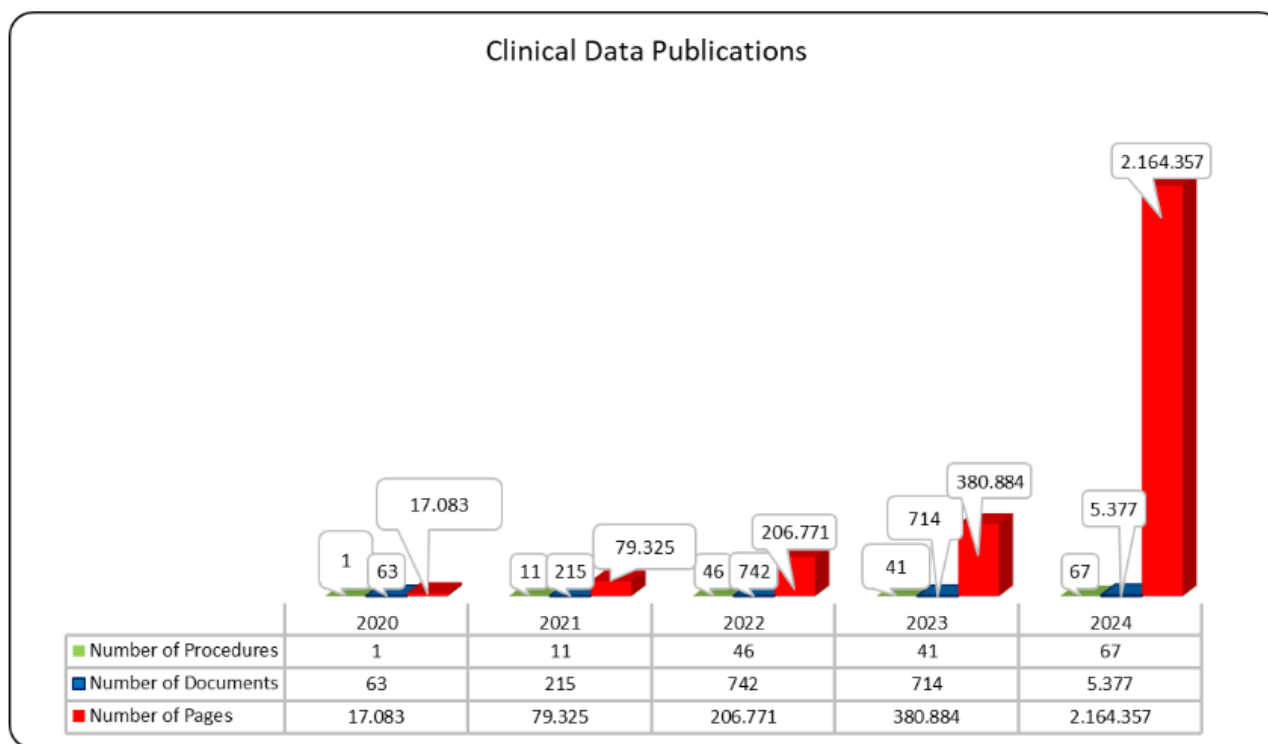
Table 2: Minimum items recommended by WHO for reporting summary results in trial registries

Transparency of clinical trials in the EU

- As of October 2016, the EMA publishes clinical data submitted by pharmaceutical companies to support their regulatory applications
- New applications and COVID transparency measures (current step 1)
- Clinical study reports - detailed documents prepared by trial sponsors for regulatory agencies
- Methods and results (aggregate data, with few exceptions)
- EU transparency resources: EMA Clinical Data Portal and Clinical Trials Information System (CTIS)

Increase transparency of clinical trials in the EU

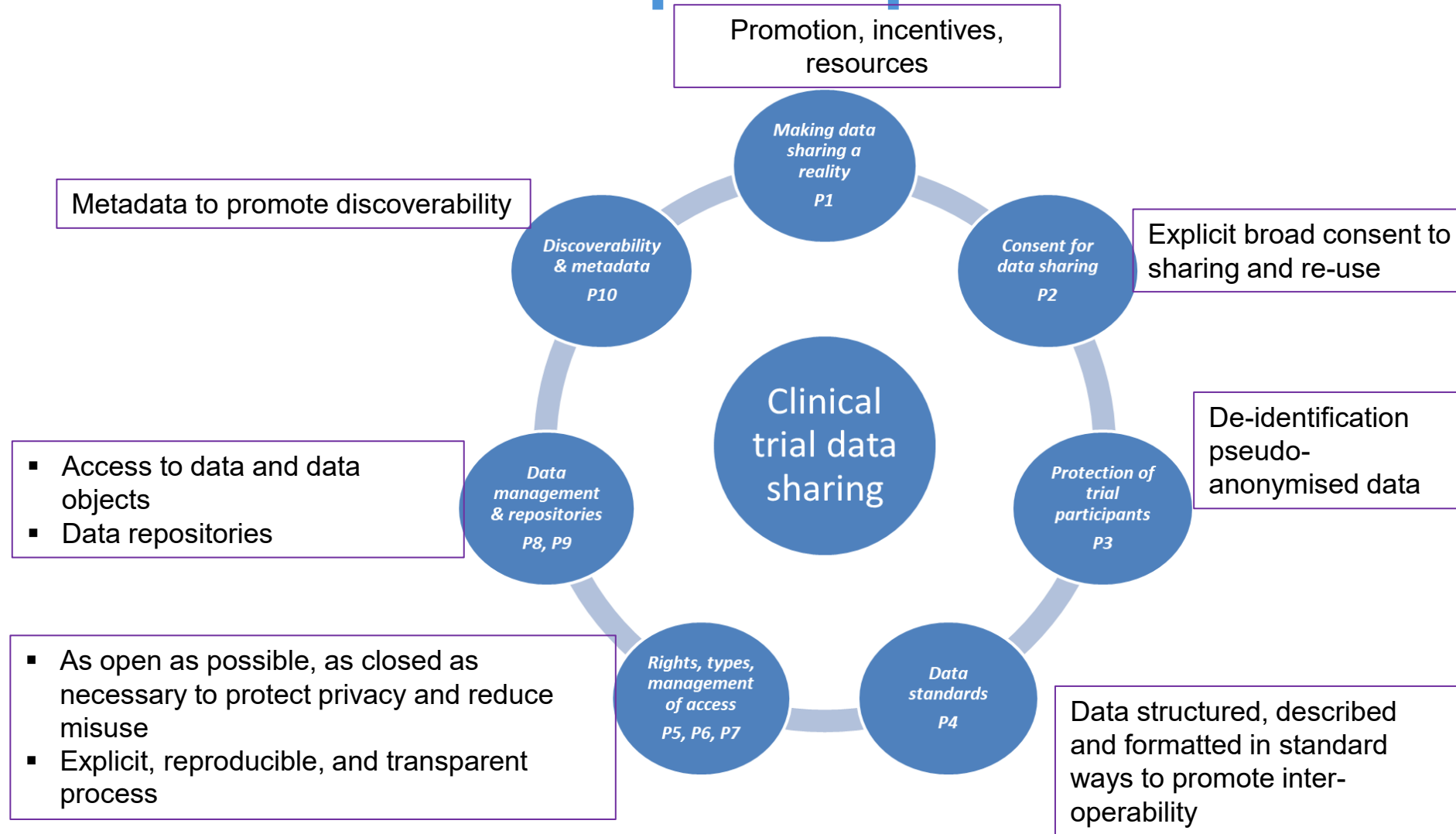
CDP Data from 2020 to Oct 2024

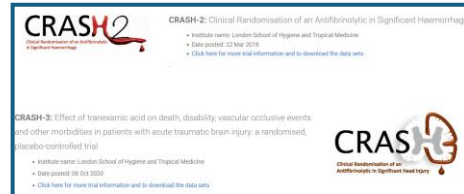


What is missing?

- Clinical study reports available case-by-case and request-based, not fully open by default
- Heavily redacted (especially “commercially confidential information”), sometimes delayed or incomplete
- Non-drug health intervention studies (weaker mandatory registration, less systematic results reporting, more fragmented oversight across countries and disciplines)
- Routinely availability of participant-level data

Access to individual participant data





[CRASH_data-1.csv](#)
[Data_Dictionary_CRASH_data.docx](#)
[Data_Dictionary_CRASH_data1.pdf](#)

Trial website:
<http://www.crash.lshtm.ac.uk>
Contact email:
CTU@lshtm.ac.uk
Contact phone:
+44(0)20 7299 4684

CRASH trial collaborators. Effect of intravenous corticosteroids on death within 14 days in 10008 adults with clinically significant head injury (MRC CRASH trial): randomised placebo-controlled trial. *Lancet* 2004; 364: 1321-28

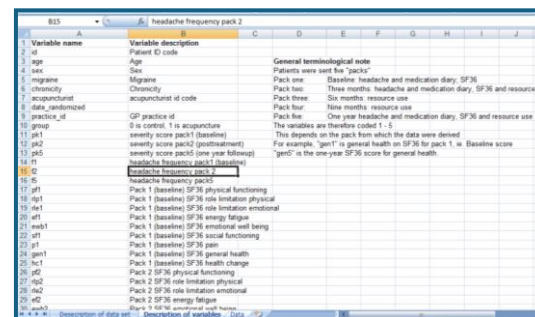
CRASH trial collaborators. Final results of MRC CRASH, a randomised placebo-controlled trial of intravenous corticosteroid in adults with head injury - outcomes at 6 months. *Lancet* May 2005 published on-line DOI: 10.1016/S0140-6736(05)66552-X

Collaborative approach...

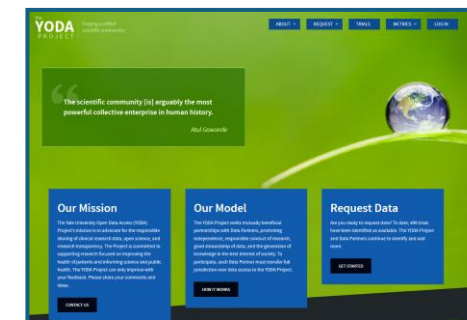


«naive» approach...

A recent Cochrane review of 26 randomised trials of acupuncture for headache concluded that, although existing evidence supports the value of acupuncture, the quality and amount of evidence are not fully convincing.⁷ The review identifies an urgent need for well planned, large scale studies to assess the effectiveness and cost effectiveness of acupuncture under "real" conditions. In 1998 the NHS National Coordinating Centre for Health Technology Assessment commissioned us to conduct such a trial (trial number ISRCTN06537534). Our aim was to assess the effectiveness of acupuncture in the treatment of headache in a primary care setting. The trial was a randomised controlled trial comparing acupuncture with a control group. The control group received a standardised treatment of paracetamol and ibuprofen. The acupuncture group received a standardised treatment of acupuncture. The primary outcome was the number of headache days in the last 4 weeks. The secondary outcome was the number of headache attacks in the last 4 weeks. The trial was conducted in 1999 and 2000. The results of the trial are presented in this paper.



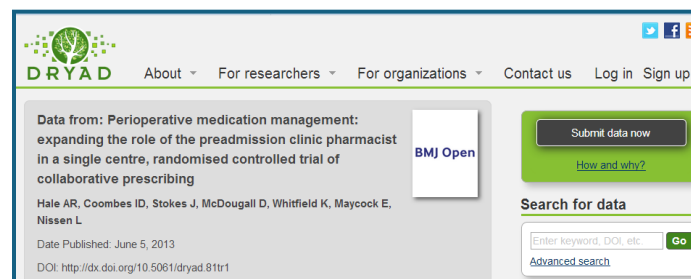
Platform-based data sharing



<https://yoda.yale.edu/>



<https://vivli.org/>



Hale BMJ Open 2013

	Guidelines for optimal and storage	Dis-identification	Data quality control	Contract for optimal and storage	Application of metadata	Application of identifiers	Flexibility of Access	Long term Preservation
Orvid								
Worldwide National Data Service								
Open								
ELIAS								
Euphonia								
EPH								
SDCT/SHIR								
NIH/NCIN/CIHR/Alberta								
NIH/NIH/NIH/NC								
Healthcare/Healthcare								
Unit								
GENOME								
Open Science Framework								
Project/Integration								
Zenodo								
NCBI								
FIN/Finland								
Genomics/Gen								
SWA/EN								
Accession/Map								
Protein								
FIN/WHO								
Genomics								
SWA/EN								
THE/INSTRUMENT								

Legend

Documented
 Not Documented
 Partially documented
 Missing or partial information available

Fig. 2 Suitability of the repository for hosting clinical study data

Banzi Trials 2019

Vickers BMJ 2004

Access to individual participant data

Everyone should do their own part:
funders

Core funders of medical
research commit to
strengthening clinical trials
worldwide



Joint statement on strengthening clinical trials 2025

25 September 2025

III. Clinical trials that are supported to be conducted to meet best-practice standards, by including the following elements in trial policies or conditions of funding:

- clinical trials are conducted in line with the WHO Guidance for Best Practices for Clinical Trials;
- proportionate engagement of communities, to be conducted throughout the trial lifecycle, as an essential component of ethical trials;
- registration and update on trial progress in a publicly available, free to access, searchable, clinical trial registry complying with WHO's international agreed standards and updating of registries to include trial results, working towards a timeframe of 12 months from primary trial completion (in line with the WHO Joint Statement on the Public Disclosure of Results);
- encourage the use of standardized data protocols where available and Core Outcome Sets;
- open-access publication of clinical trial materials (such as trial protocols and statistical analysis plans) at the earliest opportunity, and preferably through trial registries;
- timely publication of results (working towards a timeframe of 12 months from primary study completion) including reporting of outcome and adverse event data disaggregated by sex/gender and age, preferably in an open access peer-reviewed journal, with a trial registration ID and data availability statement detailing how the data underlying the publication can be accessed;
- encourage researchers during a public health emergency to rapidly and responsibly share interpretable results, including negative results, with relevant authorities for clinical guideline development and emergency use listing; and
- encourage sharing of de-identified data (or meta-data where required), complying to international data standards, in a suitable repository with a persistent identifier.

Access to individual participant data

Everyone should do their own part: medical journals



2. Data Sharing

The ICMJE's data sharing statement policy is detailed in an editorial (see [Updates and Editorials](#)).

1. As of 1 July 2018 manuscripts submitted to ICMJE journals that report the results of clinical trials must contain a data sharing statement as described below.
2. Clinical trials that begin enrolling participants on or after 1 January 2019 must include a data sharing plan in the trial's registration. The ICMJE's policy regarding trial registration is explained [above](#). If the data sharing plan changes after registration this should be reflected in the statement submitted and published with the manuscript, and updated in the registry record.

Data sharing statements must indicate the following: whether individual deidentified participant data (including data dictionaries) will be shared ("undecided" is not an acceptable answer); what data in particular will be shared; whether additional, related documents will be available (e.g., study protocol, statistical analysis plan, etc.); when the data will become available and for how long; by what access criteria data will be shared (including with whom, for what types of analyses, and by what mechanism). Illustrative examples of data sharing statements that would meet these requirements are provided in the Table.

www.icmje.org/recommendations/browse/publishing-and-editorial-issues/clinical-trial-registration.html

Taichman Annals Int Med 2017

Data and code sharing for research papers in The BMJ

The BMJ requires a data sharing statement for all research papers. Additional requirements differ depending on the type of research.

Clinical Trials submitted on or after 1 May 2024:

For reports of clinical trials submitted to *The BMJ* on or after 1 May 2024, we require that the data underlying the results be available at the time of publication. This encompasses all anonymised data on individual patients on which the analyses, results, and conclusions reported in the paper are based.

Data repositories should:

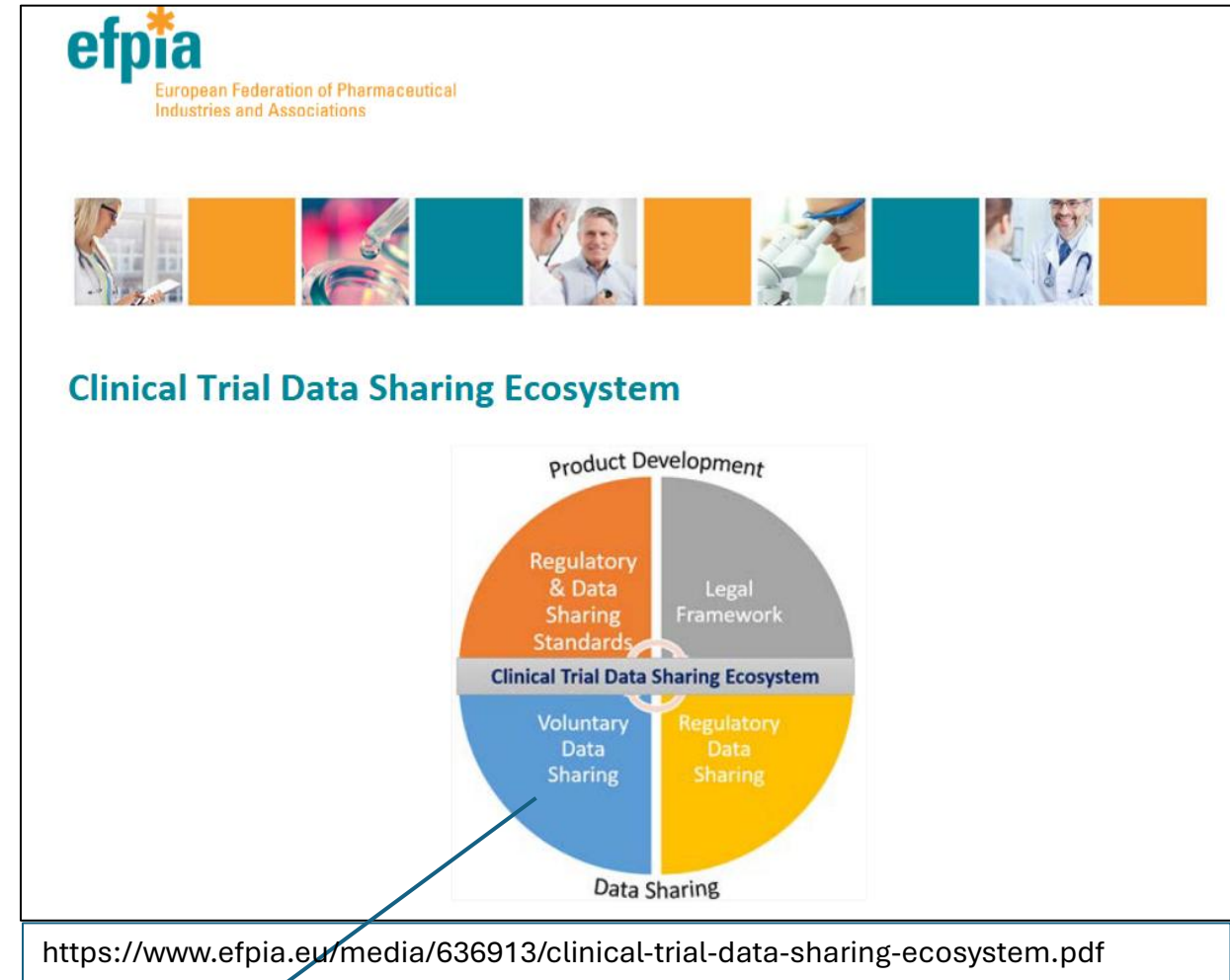
- Have open licences or allow access without unreasonable barriers.
- Provide a stable and persistent identifiers for datasets, such as a DOI.

[Dryad](#) is a suitable digital repository and allows open deposition and direct linkage by DOI from the dataset to the article and back ([see here](#) for datasets from BMJ journal articles). Other examples of suitable repositories include [Vivli](#), [Synapse](#), and the [UK Data Service](#). Authors should also investigate whether there are institutional, funder, or discipline-specific databases available for data deposition that meet the above standards.

<https://www.bmj.com/about-bmj/resources-authors/article-types/data-sharing>

Access to individual participant data

Everyone should do their own part:
industries



Access to individual participant data

**Everyone should do their own part:
researchers (especially young researchers!!!)**



SHARING AND RE-USING CLINICAL
TRIAL DATA TO MAXIMISE IMPACT



Funded by
the European Union

HORIZON-MSCA.2022-DN 101120360



...and more in general meta-researchers

OSIRIS

OPEN SCIENCE TO INCREASE REPRODUCIBILITY IN SCIENCE

„Reproducibility is crucial to progress and impact of Research and Innovation (R&I) as it confirms or corrects the outcomes of single studies, resulting in higher quality research, more reliable and implementable outcomes, and reduction of research costs.“

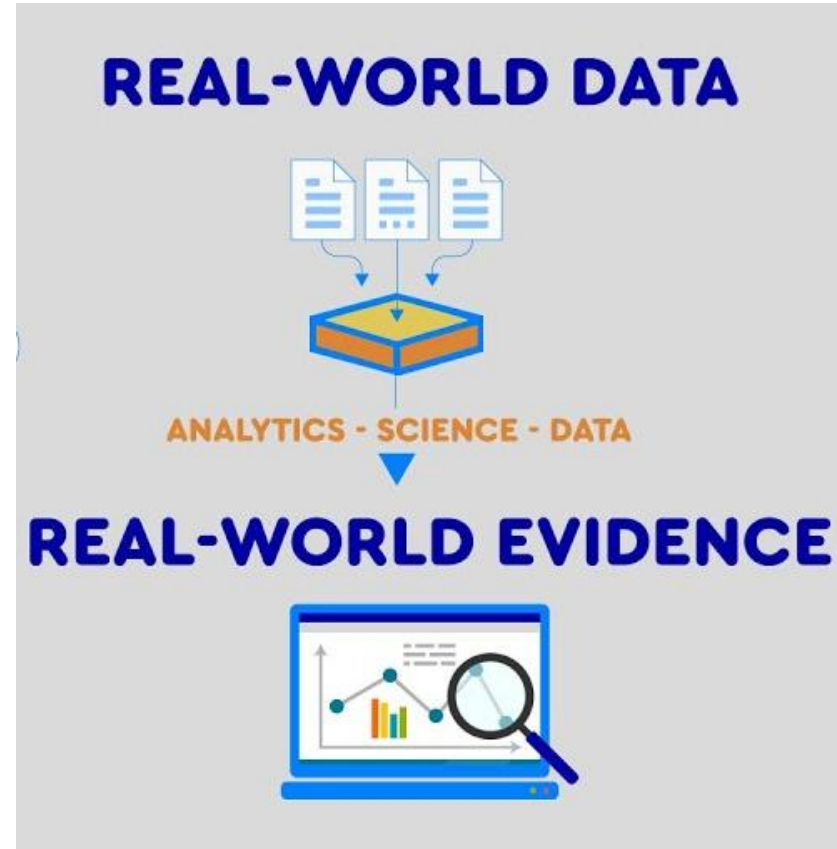
“Creating Trust in Open Science & Reproducibility through Accessibility and Transparency!”

**Beyond clinical trial data, let's make all
types of data accessible and reusable**

Routinely collected data as a complement to clinical trial data



Often insufficient to guide decision-making, as they are inherently unable to assess the impact of treatments in 'real-world clinical practice'



Growing interest in the development of methods capable of generating reliable evidence on the impact of care pathways using data collected in the 'real world'

Using routinely collected data for research purposes: challenges and mitigation strategies

Sabine Hoffmann,¹ Tim Morris,² Moritz Herrmann,^{3,4} Georg Heinze,⁵ Laure Wynants,^{6,7,8} Ben Van Calster,^{7,8} Bernd Bischl,^{1,4} Matthias Schmid,⁹ Pamela A Shaw,¹⁰ Tim Mathes,^{11,12} Florian Naudet,^{13,14} Frank E Harrell,¹⁵ Mona Niethammer,¹ Samuel Muli,¹⁶ Sebastian Zimmer,¹⁷ Farhad Bakhtiary,¹⁸ David Leistner,¹⁹ P Christian Schulze,²⁰ Daniel Sedding,²¹ Philipp Lurz,²² Tienush Rassaf,²³ Malte Kelm,²⁴ Stephan Baldus,²⁵ Johann Bauersachs,²⁶ Georg Nickenig,¹⁷ Holger Thiele,²⁷ Enzo Lüsebrink¹⁷

No evidence for a research question may be preferable to evidence from a low quality routinely collected data study, as results may otherwise be misleading and overinterpreted

(I bet many decision-makers would disagree...)

SUMMARY POINTS

Routinely collected data, which are increasingly used for research purposes, offer significant opportunities in terms of time and cost of data collection, but researchers often lack knowledge of how the data were generated and control over how they were collected

Common challenges in the analysis of routinely collected data include representativeness, time point alignment, data quality, interventions and tests not being random, and analytical variability due to a multiplicity of possible analysis strategies

The flexibility of AI algorithms can make them highly vulnerable to the biases arising in the analysis of routinely collected data

Without rigorous internal and external validation, and in the absence of clinical and methodological expertise, these AI tools may do more harm than good when analysing routinely collected datasets

Lack of data integrity may mean that prospectively collecting research data will often be better, as no evidence for a research question may be preferable to evidence from a low quality study that may be misleading and over-interpreted

Critical research areas – emergency medicine

November 10, 2025



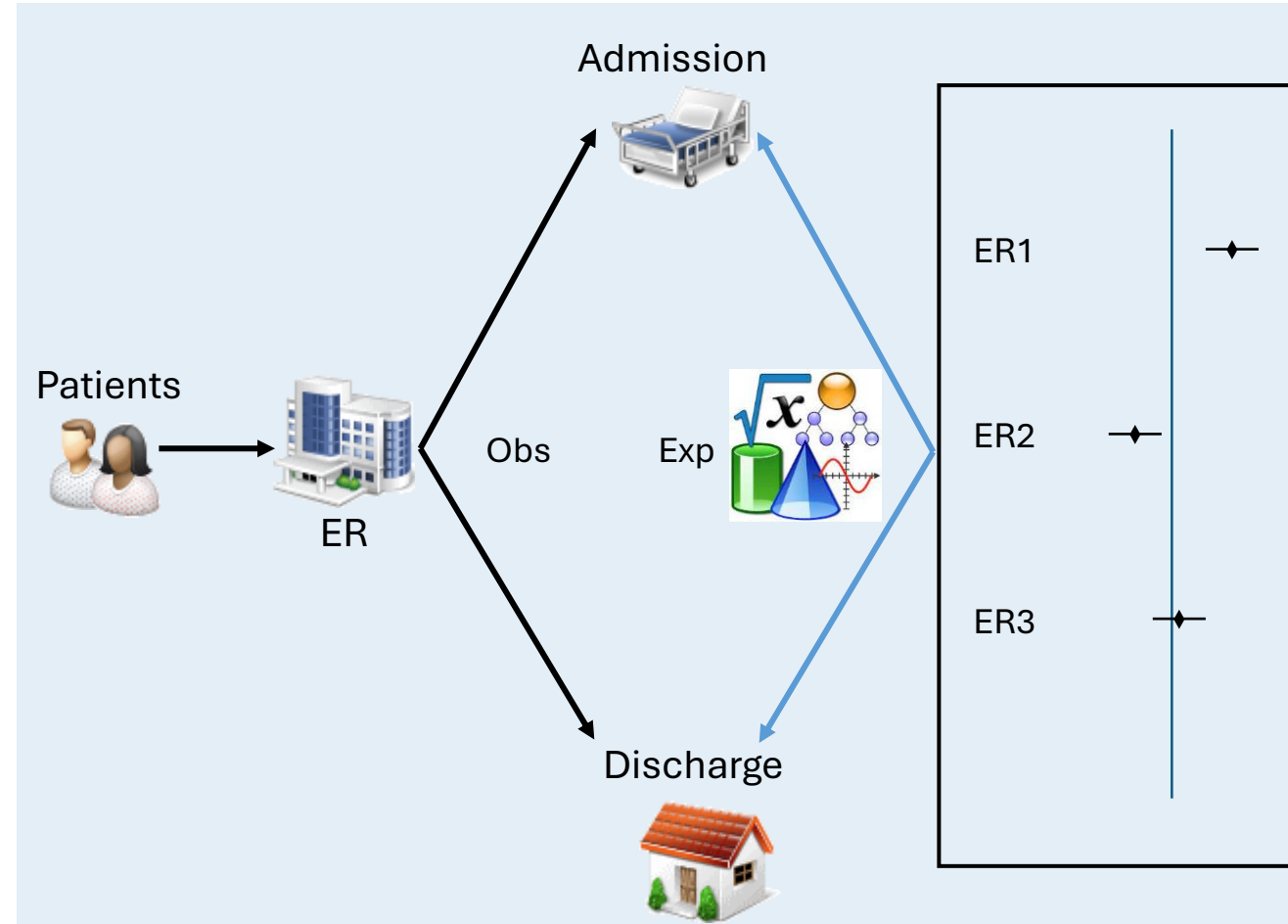
Critical research areas – emergency medicine

Observational study in the emergency department to collect **clinical and organizational variables** influencing the likelihood of hospital admission

- Average time for obtaining and managing consent: **15 minutes**
- Average time for completing the data collection form: **20 minutes**
- Average time emergency physicians currently spend per patient: **42 minutes** (clinical + administrative activities)



Conducting this study would effectively **double the time spent per patient**



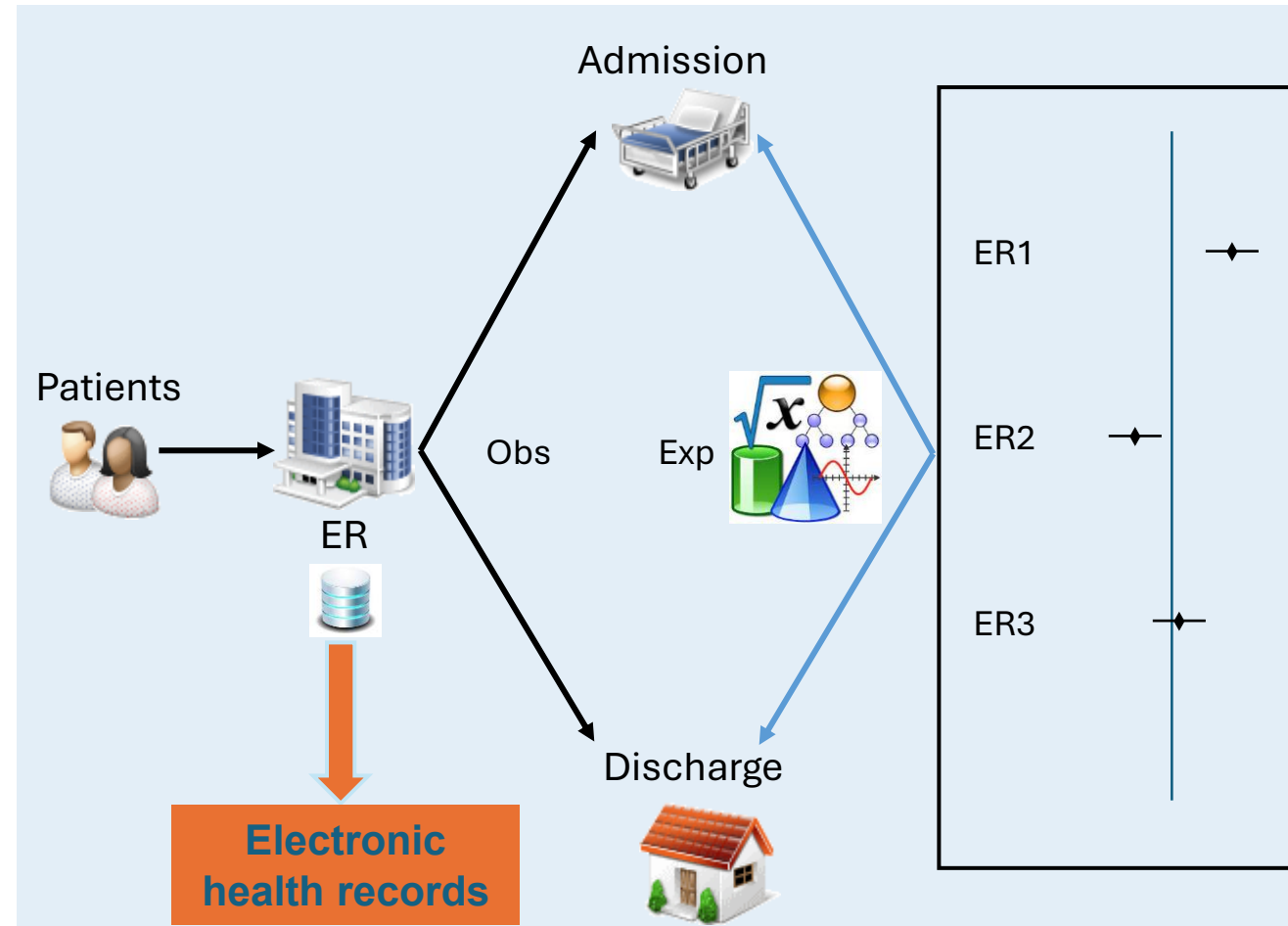


Critical research areas – emergency medicine

Clinical and organizational variables...the data we need already exist!

We just need to “**only**” find a way to:

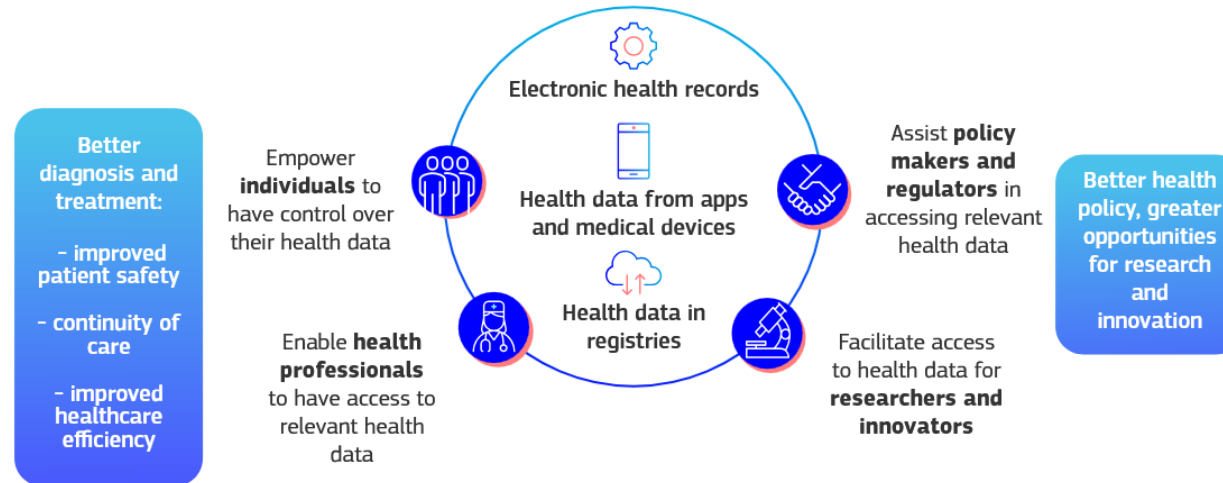
- locate them
- understand them
- use them in compliance with regulations
- integrate them
- make them available for research (ours and others’)



Reuse of health data for research purposes

“What a time to be alive!”

Upcoming implementation of the regulation on the European Health Data Space (EHDS), a transformative step in the governance of health data in Europe



March 2025:
The EHDS Regulation entered into force (transition period started)

March 2027: Deadline for the Commission to adopt several key implementing acts, for the regulation operationalisation

March 2029:
Key parts of the EHDS Regulation will enter into application, Rules on secondary use will also start to apply for most data categories

To be continued.....

Final remarks

- Transparency of clinical research is an increasingly important issue for regulators and health decision makers worldwide as it improves evidence-based decision-making, patient safety, and public trust in medicine and research
- We should look at the entire ecosystem of evidence alongside the entire ecosystem of decision-makers
- The research community must remain vigilant and continuously develop robust, appropriate methods to ensure that evidence is reliable, transparent, and fit for decision-making

Thank you!

rita.banzi@marionegri.it

