EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

Changing the discoverability and data quality of EU data ecosystem

HMA/EMA Big Data Stakeholder Forum 2023

4 December 2023
Presented by Susana Perez-Gutthann (ENCePP)
Acknowledgements to Ana Cochino, Catherine Cohet (EMA)

Disclosures & Perspectives

EUROPEAN MEDICINES AGENCY

Co-Chair of the Steering Committee of the European Network of Centers for Pharmacoepidemiology and Pharmacovigilance (**ENCePP**) and past Co-Chair of the working group on **Research Standards**

Past President and long-time service as officer of International Society for Pharmacoepidemiology (**ISPE**), including chair public policy committee (**Good Pharmacoepidemiology Practice**)

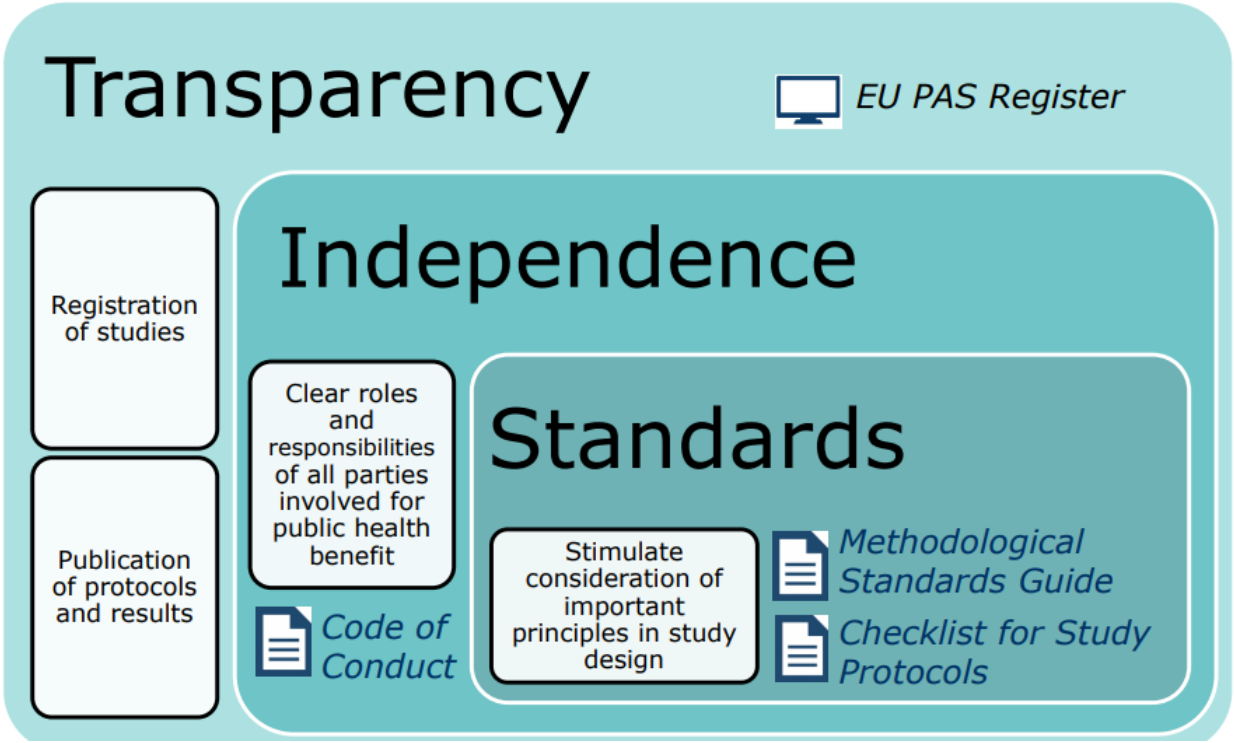
- Employed at RTI, independent non-profit research institute working for government and other institutions. My work focuses on regulatory grade RWE, mostly funded by pharma.
- Board Chair of the SIGMA Consortium, hub for regulatory RWE studies
- Past employment 1990-2007 R&D pharma epidemiology

The **European Network of Centres for Pharmacoepidemiology & Pharmacovigilance** is an **EMA-led** network of research centers to strengthen the

- Methods
- Study conduct
- Transparency (registration)

Of post-authorization studies in the EU

Pillars & Tools

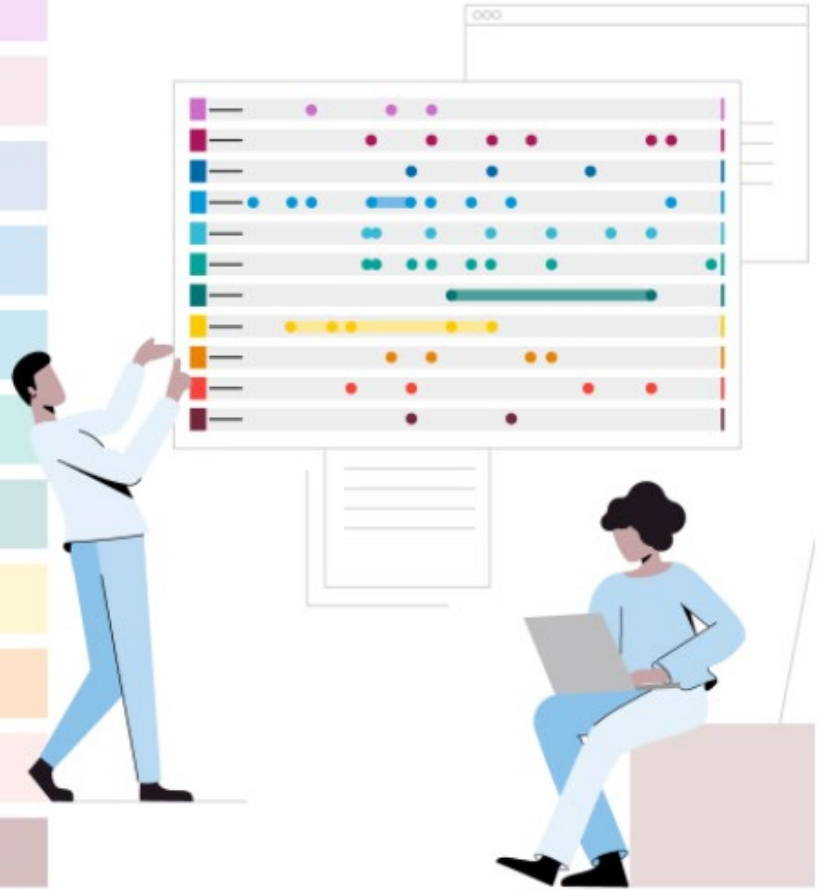


[ENCePP Home Page]

HMA-EMA joint Big Data Task Force

EUROPEAN MEDICINES AGENCY

- DARWIN EU
- Data quality & representativeness
- Data discoverability
- EU Network skills
- EU Network processes
- Network capability to analyse
- Delivery of expert advice
- Governance framework
- International initiatives
- EU stakeholder implementation forum
- Veterinary recommendations




PERSPECTIVES

PERSPECTIVE

Real-World Evidence in EU Medicines Regulation: Enabling Use and Establishing Value

We outline our vision that by 2025 the use of real-world evidence will have been enabled and the value will have been established

Arlett et al., Clin Pharmacol Ther. 2022 Jan

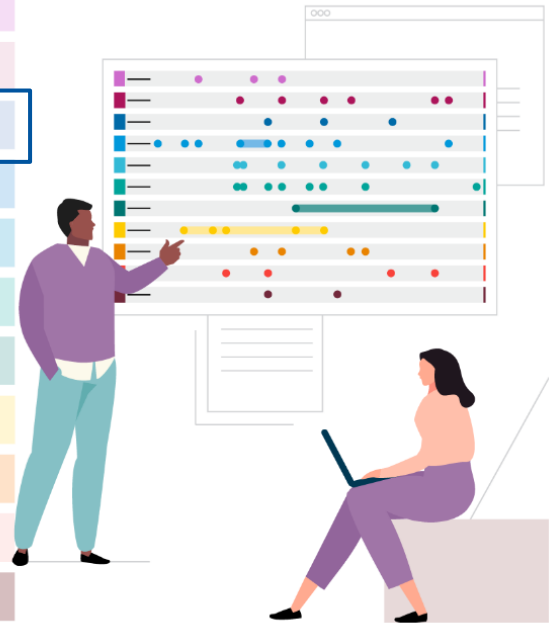
Big data | European Medicines Agency (europa.eu)

Data discoverability

Recommendation 3:
Enable data discoverability

- Identify **Key Metadata for regulatory decision-making** on the choice of data source
- **Strengthen the current ENCePP** Resource Database

DARWIN EU
Data quality & representativeness
Data discoverability
EU Network skills
EU Network processes
Network capability to analyse
Delivery of expert advice
Governance framework
International initiatives
Stakeholder engagement
Veterinary recommendations



Catalogue of data sources

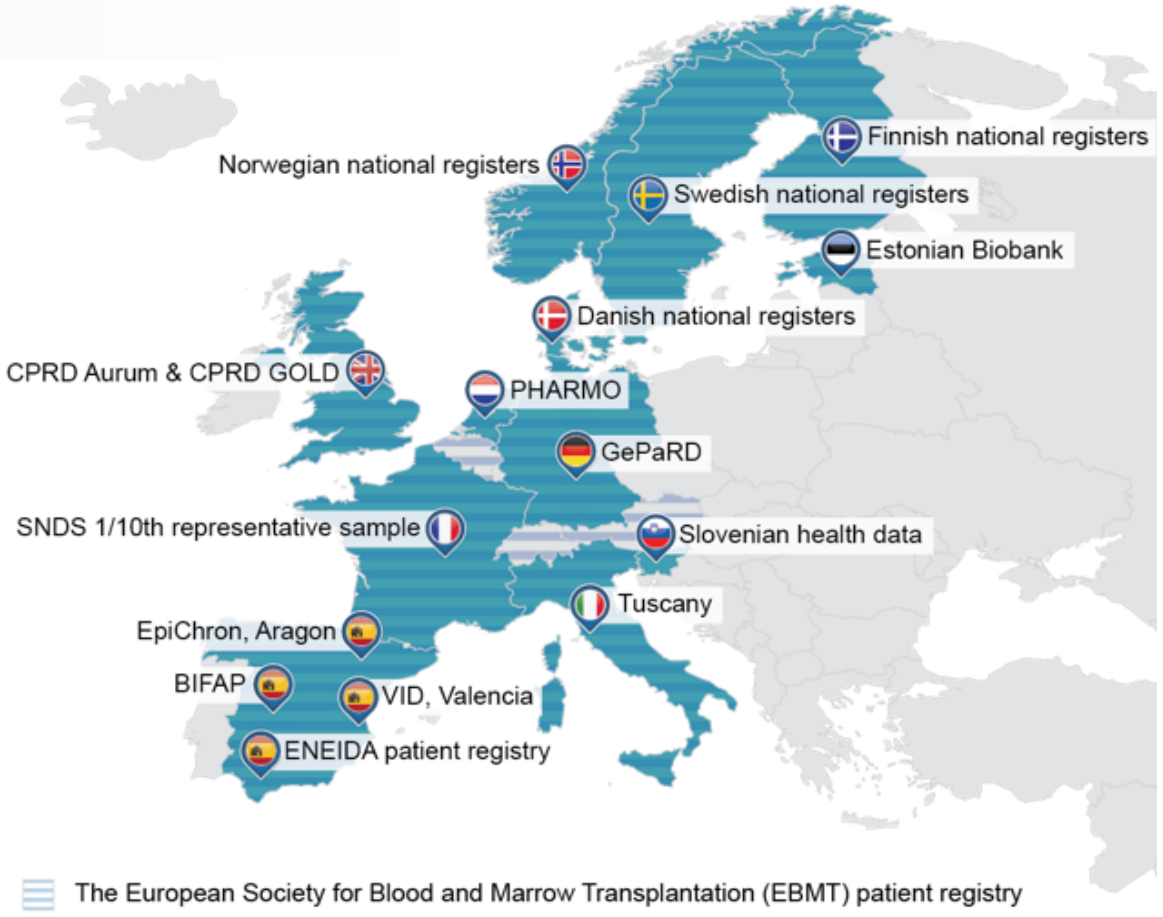
Catalogue of non-interventional studies

Metadata Pilot - MINERVA Consortium & EMA

EUROPEAN MEDICINES AGENCY

- **15** RW population data sources and patient registries in **12** countries
- **18** research centers in **12** countries

RTI-HS, Barcelona, Spain	Σ	BIPS, Bremen, Germany	Σ
UMCU, Netherlands	Σ	CPE KI, Stockholm, Sweden	
UU, Netherlands	Σ	LSHTM, London, UK	
ARS, Toscana (ARS), Italy	Σ	AEMPS, Madrid, Spain	
UMCG, Netherlands		FISABIO, Valencia, Spain	
BPE, France	Σ	IACS, Zaragoza, Spain	
DCE-AU, Denmark	Σ	UL FFA, Ljubljana, Slovenia	
PHARMO, Netherlands	Σ	GETECCU, Santiago de Compostela, Spain	



6

MINERVA Conclusions & Lessons Learned

EUROPEAN MEDICINES AGENCY

EU PAS Register Study EUPAS39322 LINK
Protocol, reports, recommendations, list of metadata

The MINERVA pilot showed the value of piloting major metadata catalogue processes and a need for data curation requiring **epidemiologic** and **data prompt** knowledge

Setting up and maintaining an operational metadata catalogue on real-world data sources requires **substantial effort** to implement FAIR principles, adhere to data protection rules, and effectively support discoverability of data sources and reproducibility of studies in Europe.

Good Practice Guide - update

- The **Good Practice Guide (GPG)** aims to **provide recommendations** for the use of the EMA-HMA catalogue to **identify real-world data sources suitable for specific research questions** and to **assess the suitability** of data sources proposed to be used in a study protocol or referred to in a study report.

- It also provides a **detailed description of all the metadata elements** as envisaged to be used in the catalogue. List of metadata for Real World Data catalogues

- The GPG also **guides the user for the insertion and maintenance** of data in the catalogue.

Public consultation: Contributions received from the regulatory network, data holders, researchers and research networks, industry, medical associations, etc.

Amended scope: GPG document will be split into two parts:

1. Good practice guide: Use of Catalogue to assess the suitability of data sources→ **Q1-Q2 2024**
2. User Guides → **at Go-live of catalogues**

EMA-HMA Catalogues of data sources & non-interventional studies

EUROPEAN MEDICINES AGENCY

The **EMA-HMA Catalogues of data sources and non-interventional studies** will describe **real-world data sources and studies** through a set of collected **metadata** to help pharmaceutical companies and researchers **identify and use** such data when investigating the use, safety and effectiveness of medicines.

Catalogue of data sources

will replace and enhance the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (**ENCePP**) Resources Database

Catalogue of studies

will enhance the European Union electronic register of post-authorisation studies (**EU PAS Register®**)

- **Efficient and user-friendly platform** for researchers, regulators, and pharmaceutical companies
- **Centralised and enhanced resources** that contribute to the transparency of observational research
- Promotion of good practices aligning with **'FAIR' data principles** for **F**indable, **A**ccessible, **I**nteroperable, and **R**eusable data
- Facilitation of **search and evaluation of data sources and studies** related to medicines, ultimately supporting evidence-based decision-making

Integration with other catalogues, EHDS and similar initiatives to be developed in coming years

Search all content

My search term

Search

Filter options

- Data source
- Institution
- Network
- Study

Country

Germany

Regions (geographical regions that the data source covers)

Select Value

Data source type

Hospital inpatient records

Data Holder

Select Value

Apply Clear

Results (3)

Sort by Newest first

Data source Germany Hospital inpatient records

Export results

Hepatitis Delta International Network (HDIN) - Patient Registry

Austria Belgium Brazil Germany Greece Italy Mongolia Pakistan Romania Turkey United States

First published: 19/09/2023 **Last updated:** 08/11/2023

Data source | Hospital inpatient records | Hospital outpatient visit records | Primary care medical records

Hospital Medical Records Database DE

Austria Germany

First published: 31/10/2023 **Last updated:** 31/10/2023

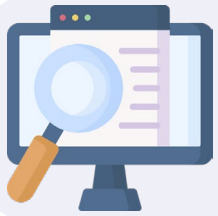
Data source | Hospital inpatient records

Deutsche Leberstiftung (German Liver Foundation)

Azerbaijan Belgium Brazil Georgia Germany Greece Moldova, Republic of Mongolia Pakistan Spain Sweden Turkey United States Viet Nam

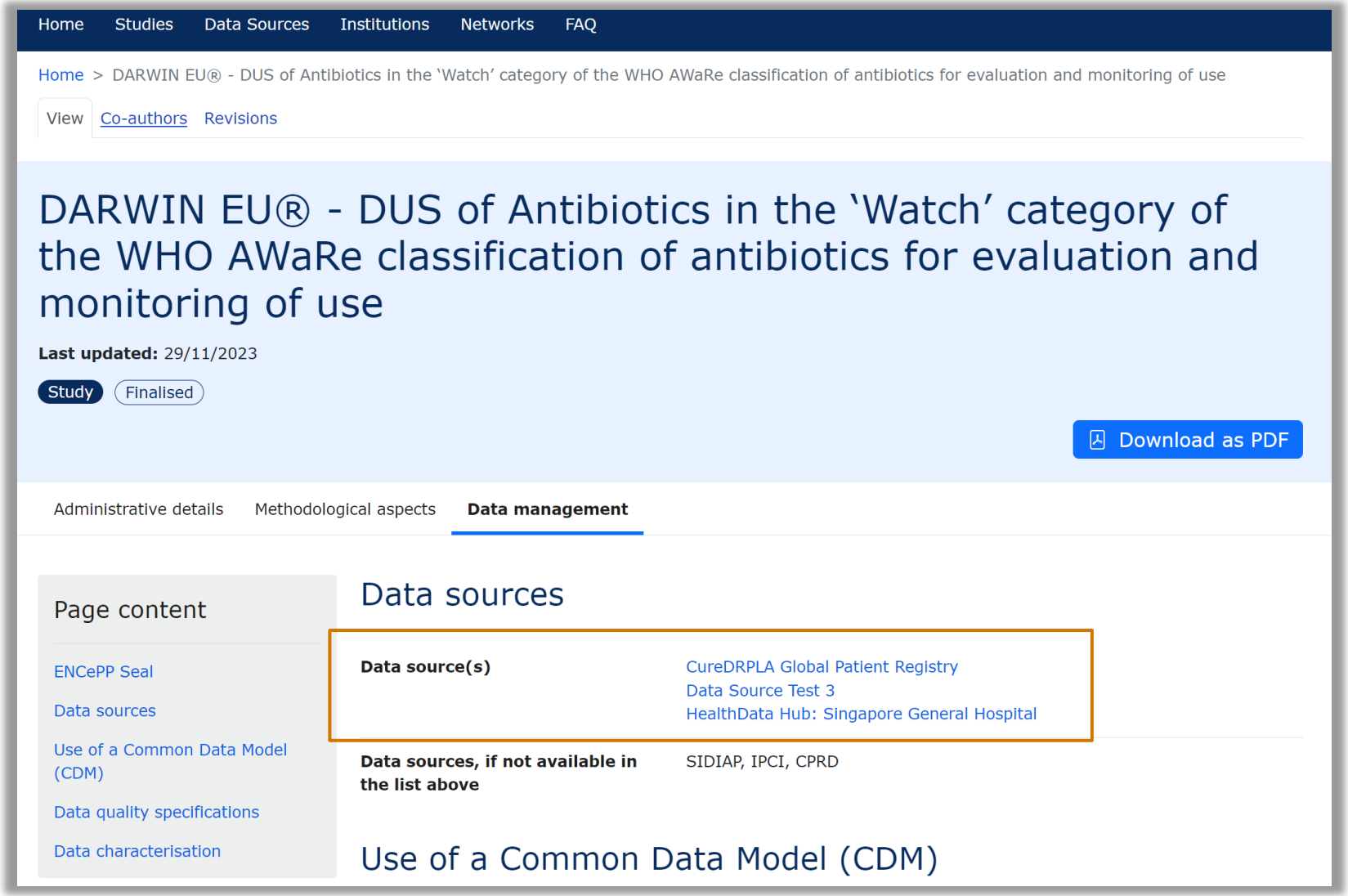
First published: 19/09/2023 **Last updated:** 30/10/2023

Data source | Hospital inpatient records | Hospital outpatient visit records | Primary care medical records



Enhanced search & export functionalities possibility to filter, sort and export search results and records

Link between **data sources and associated studies**



Home | Studies | Data Sources | Institutions | Networks | FAQ

Home > Add content > Data source

Add data source

Please complete the questionnaire to register your data source in the EMA-HMA catalogue of data sources. Mandatory fields are marked with an asterisk (*).

The information provided in the questionnaire needs to be kept up-to-date by the editor of this entry, and this is not the responsibility of the EMA.

The questionnaire's 16 questions are divided into 4 steps: 1. Administrative Details, 2. Data Elements Collected, 3.
and Management. A sample questionnaire for offline review only, can be downloaded using the following link: (FILE

You agreed with the terms and conditions when you joined the EMA-HMA Catalogues.

STEP 1 Administrative details | STEP 2 Data elements collected | STEP 3 Quantitative descriptors | STEP 4 Data flows and management

Administrative details

Name *

The name of the data source, as used in European projects, must be provided. If the database is widely known by several names, t
can be provided in this field, separated by a '/' sign. Where the data source has been known by different names in the past, these c
provided, using parenthesis with the note 'formerly known as'. Where the name of the data source is in a local language, the Engl
translation should also be provided, using parentheses.

Welcome to your dashboard

+ Add content

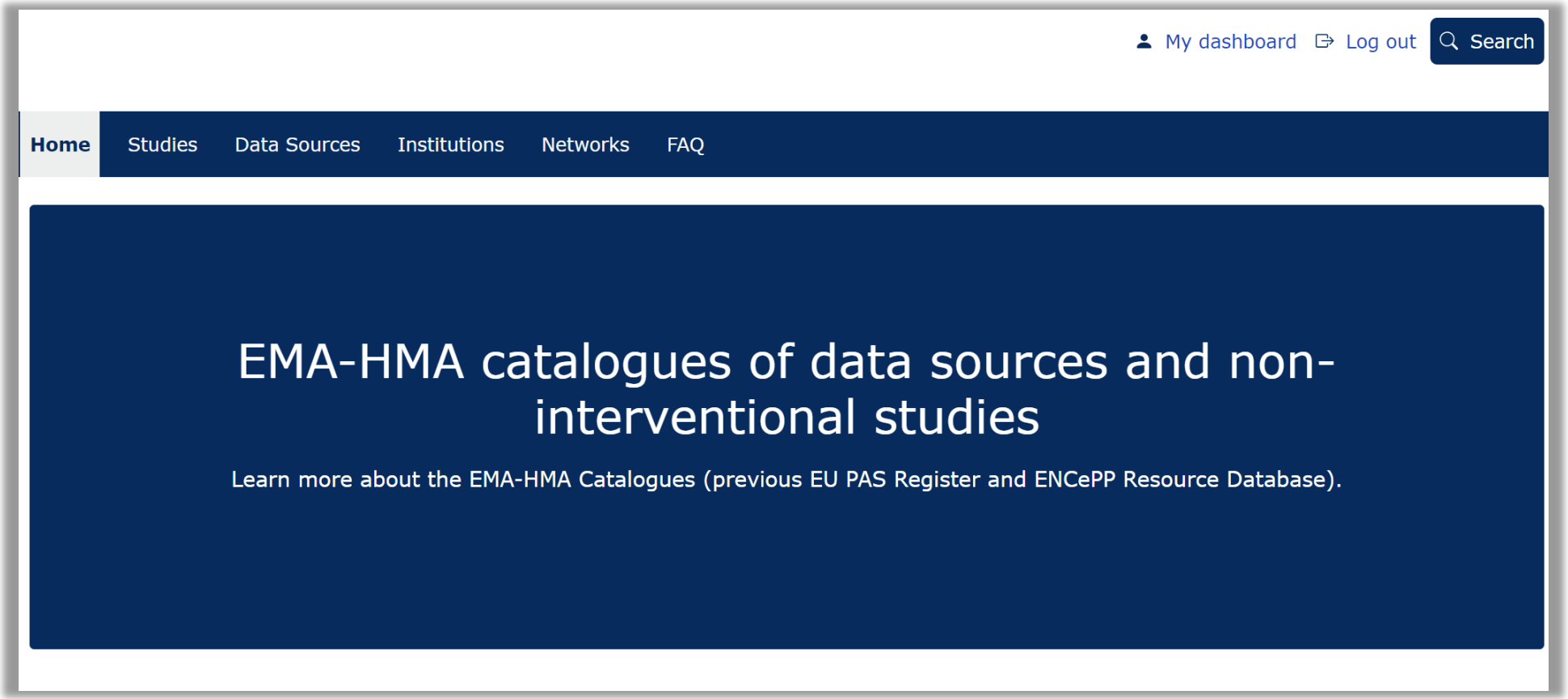
My collaboration requests

TYPE	TITLE	INVITED BY	UPDATED	STATUS	
Study	EXPLORING THE BIDIRECTIONAL RELATIONSHIP BETWEEN DATABASE MARKERS OF ASTHMA TREATMENT ADHERENCE AND ASTHMA-RELATED OUTCOMES		Thu, 19/10/2023 - 10:06	Accepted	Decline
Data source	Biologika in der Kinderrheumatologie		Thu, 19/10/2023 - 10:54	Accepted	Decline

TYPE	TITLE	UPDATED	PUBLICATION STATUS	MODERATION STATE	
Data source	Translational Research in Europe - Assessment and Treatment of Neuromuscular Diseases	26/10/2023 - 14:58	Published	Draft	Edit
Network	European Forum for Primary Care	26/10/2023 - 12:01	Published	Published	Edit

New platform enables users to **easily access and manage content** as well as **collaborate on editing** content

Integration with EMA website content: studies will be visible in the **relevant medicines overview page,** on the EMA website connection to summary of RMP, EPAR, PI*



**This feature will be released after go-live in a second phase.*

Take home messages

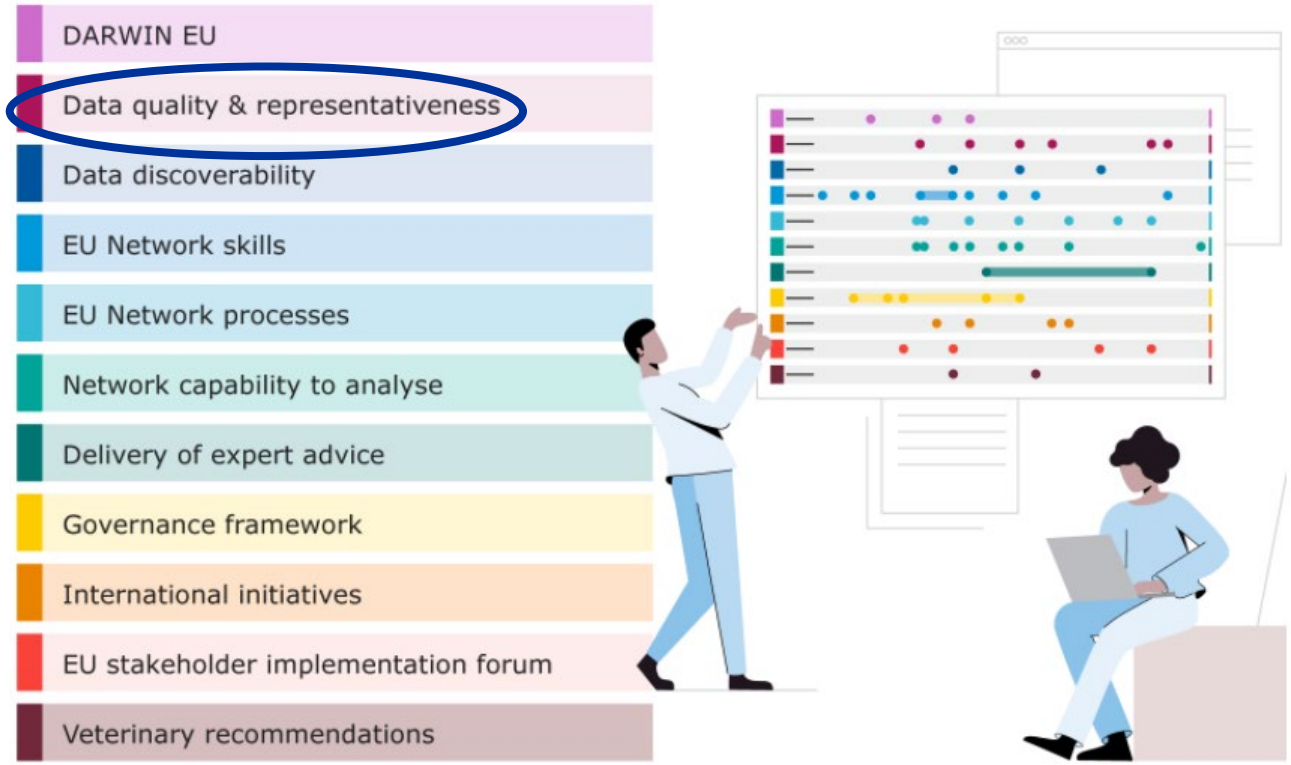
The EMA-HMA Catalogues will:

Go-Live in early 2024

Be publicly available

Data sources and studies integrated

Data Quality & Representativeness



Data Quality Framework for EU medicines regulation

Draft: consultation closed

First published: 10/10/2022
Consultation dates: 10/10/2022 to 18/11/2022

Data quality is defined as fitness for purpose for users' needs in relation to health research, policy making, and regulation and that the data reflect the reality, which they aim to represent.

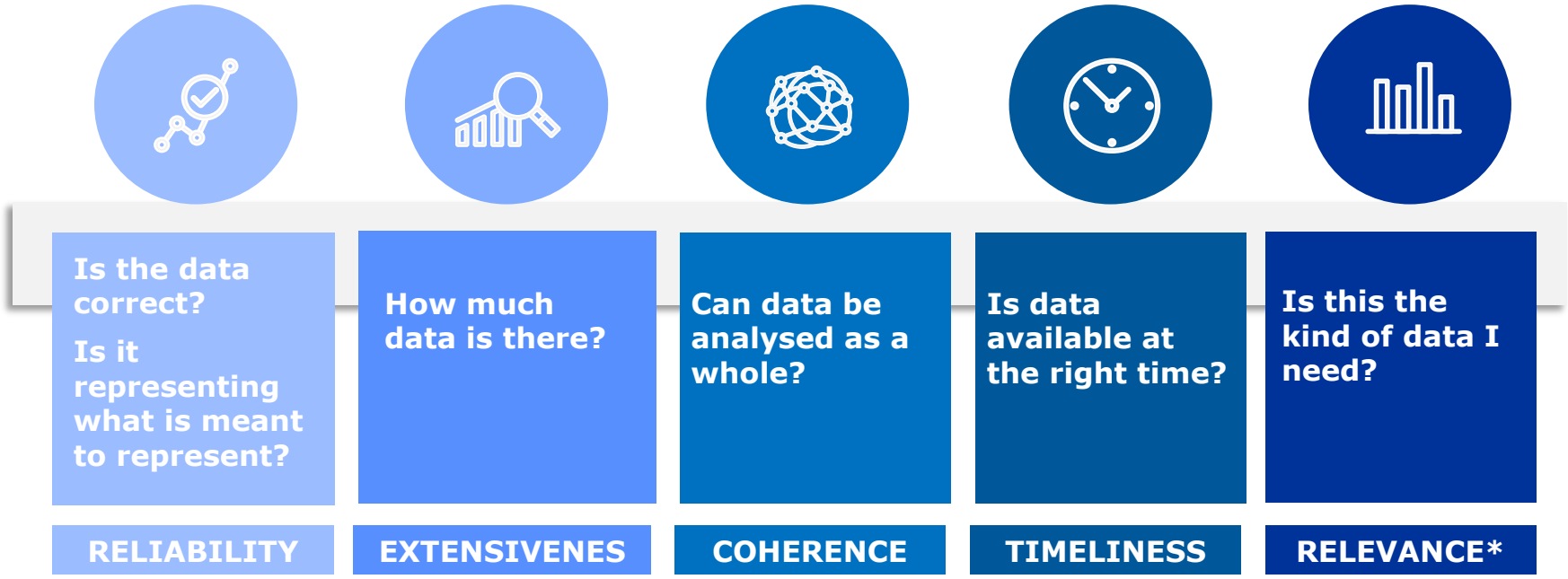
[Big data | European Medicines Agency (europa.eu)]

Data Quality Framework

EUROPEAN MEDICINES AGENCY

- Joint sponsorship of **Big Data Steering Group** & **Methodology Working Party** in close **collaboration with TEHDAS Joint Action** (Towards the European Health Data Space)
- Describes concepts **applicable to all data sources used in medicine regulation** across EU regulatory network

Main dimensions



*Relevance should be considered as the suitability to a research question

Data Quality Framework - **Real-World Data application**

EUROPEAN MEDICINES AGENCY

2021 - 2022

Landscape analysis and stakeholder consultations

- extensive landscape analysis has been produced in preparation of the work on DQF for EU regulation
- A dedicated stakeholder workshop (April 2023)

Public consultation DQF for EU medicine regulation data

- Dedicated live webinar in November 2023 introducing the document
- Wide consultation with all stakeholder groups, 400+ comments
- **Document adopted by BDSG (Big Data Steering Group), MWP (Methodology Working Party) and CHMP in 2023 (publication upcoming).**

2023 onwards

Ongoing activities:

- Preparation of draft **RWD deep dive chapter** based on the Data Quality Framework concepts and building on the input from DQF-RWD workshop

"**Opportunities for international convergence** across RWE and RWD frameworks should be considered"

"**Implications for data holders** and data submitters should be well articulated"

"The **heterogeneity of EU databases** needs to be considered when describing extensiveness"

"In relation with some research questions, it might be sufficient to **request descriptive documentation** on data reliability and relevance, as opposed to validation/verification"

Next steps:

- Public consultation of the initial draft to follow (date TBC)

The importance of methods

RWD are used to describe, predict, and for counterfactual prediction (causal inference)

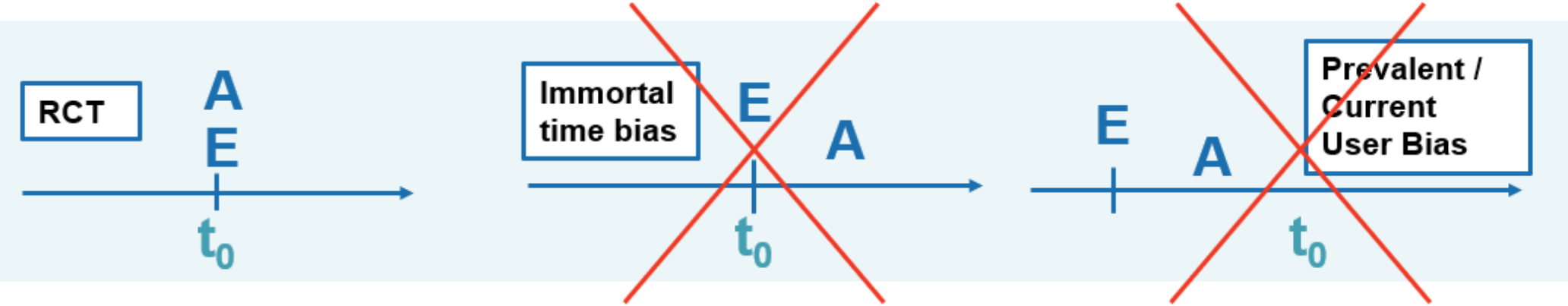
To identify a causal effect, we need data and assumptions external to the data:

Identifiability Assumptions	RCT	RWE studies
Exchangeability ***No unmeasured confounding*** The conditional probability of receiving every value of treatment, though not decided by the investigators, depends only on the measured covariates	Losses to follow-up do not happen at random	We miss baseline confounders Losses to follow-up do not happen at random
Positivity ***Exposure at all levels of the confounders*** The conditional probability of receiving every value of treatment is greater than zero	N.A.: Artificial assignment of treatment guarantees it	Data are sparse or there are too many strata
Consistency The exposures under comparison correspond to well-defined interventions aligned with the versions of treatment in the data	Protocol does not specify accurately the experimental intervention, or researchers do not follow it	Intervention is not well-defined, or data do not differentiate multiple versions of the exposure (e.g., prevalent users)

Hernán MA, Robins JM (2020). Causal Inference. Boca Raton: Chapman & Hall/CRC. https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/

Data available, confounding and positivity are under control…:

- Problems derived from the lack of synchronization in time of eligibility, treatment assignment, and time zero:
  - Time of eligibility (E): point in time when patients meet the eligibility criteria
  - Treatment assignment (A): point in time when patients are classified into exposure groups
  - Time zero (T0): point in time when follow-up starts
- Not a problem in RCT, but source of important biases in RWD studies



Garcia de Albeniz X Beyond Controlling for Confounding: Design Strategies to Avoid Selection Bias and Improve Efficiency in Observational Studies. https://www.rtihs.org/sites/default/files/Webinar_Beyond_Controlling_for_Confounding.pdf

ENCePP, a committed research community collaborating in Europe and beyond to

1- Conduct research on a foundation of principles

- Work with fit for purpose data
- Address confounding
- Align time zero, eligibility and treatment assignment

2- Collaborate with all stakeholders in recognition of the complexity and importance of understanding observational methods

3- Share lessons learned

And to support Big Data initiatives

THANK YOU VERY MUCH FOR YOUR ATTENTION

More information

For more information about the new system, you may want to refer to the product "**Real World Metadata Catalogues (RWMC)**" included in the EMA quarterly system demos listed below:

- Demo: data source form (Q2 2023)
- Demo: studies form and user dashboard (Q3 2023)
- List of metadata elements
- Catalogues good practice guide (draft version)

For questions related to the new catalogues: metadata@ema.europa.eu

Upcoming event

➢ Quarterly system demo - Q4 2023 | European Medicines Agency (europa.eu)
on 19 December