# Update on the RWD sources and RWD studies catalogues

ENCePP Plenary – 30th November 2022

Presented by Katerina-Christina Deli
Healthcare Data Workstream – Data Analytics and Methods Task Force

# Outline

1. Priority recommendation: data discoverability

2. Achievements and outlook for 2023 and beyond

3. Metadata list & Good Practice Guide – overview & comments from public consultation

4. Status update on development of the catalogues

5. EU Metadata catalogue & ENCePP website mock-ups

# Reminder – data discoverability

## Goal of setting up a catalogue of data sources and a catalogue of studies
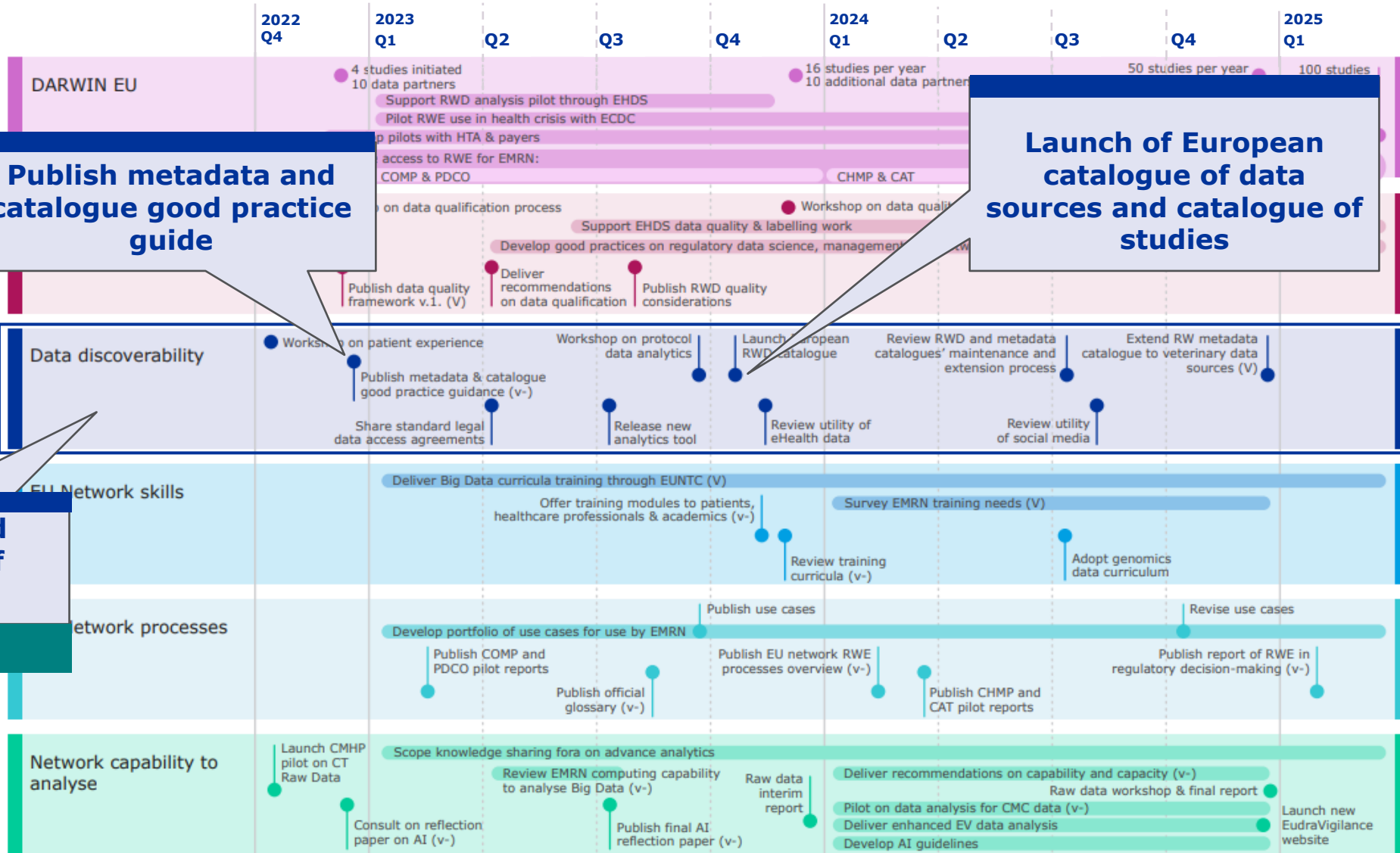
- Facilitate the **discoverability of data sources** to generate adequate evidence for regulatory purpose

  - initial identification of data sources suitable to investigate a specific research question, which could be further explored as to the accessibility and availability of the required data

- Support the **assessment of study protocols and study results** by providing quick access to information on the suitability of data source(s) proposed to be used in the study protocol or referred to in the study report

Metadata for Real World data sources and studies

# Achievements and outlook for 2023 and beyond

EUROPEAN MEDICINES AGENCY

| 2021 | 2022 | 2023 onwards |
|---|---|---|

**MINERVA pilot study**

Conducted by a consortium including 16 organisations from 9 countries

Definition of first draft of metadata list

Metadata and practices of multiple key regulators and organisations were considered as potential sources of information

Stakeholder workshop April 2021

Options for the process to collect and update the metadata

Catalogue proof of concept tool for accessing, searching, and visualising the metadata collected

---

Surveys consulting on content and functionalities of the catalogues within ENCePP community, industry, international regulators

Webinar involving TEHDAS, NCA representatives (via EUNDB, BDSG) to consult on RW metadata list

Agreement on RW metadata list for regulatory purposes

**Publication of metadata list**

Development of EU metadata catalogue started

Data collection from different data sources started

**Public consultation of good practice guide on metadata collection**

---

**Publication of good practice guide on metadata collection**

Migration of data from ENCePP and data enrichment

Data collection from different data sources (DARWIN EU®, collaboration with other catalogue initiatives) to be continued

**Launch new EU metadata catalogue**

Maintenance of real-world metadata catalogue

# Real-world metadata list and good practice guide

## Metadata list for real-world data

➢ Describes **real-world data sources and studies** to help pharmaceutical companies and researchers to **identify** and **use** such data when investigating the use, safety and effectiveness of medicines

➢ Will **feed** into two future EU catalogues:

- **Catalogue on real-word data sources** which is currently being built and will replace [European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) catalogue](#) in late 2023

- **Catalogue on real-world data studies** which will cover studies performed on the data sources, enhancing and replacing the European Union electronic register of post-authorisation studies (EU PAS Register)

## Good Practice Guide for the use of real-world metadata

In anticipation of the real-world data sources catalogue launch in 2023

➢ Provides **recommendations** to a variety of stakeholders for the **use** of the catalogue to **identify real-world data sources** for **assessing** the **suitability** of data sources for specific studies

➢ Describes the **metadata elements** that are envisaged to be used in the future EU catalogue of real-world data sources

*It's a wrap*

**Public consultation**
Good practice guide
for the use of the real-world metadata

Send comments by 16 November 2022

**EMA**
EUROPEAN MEDICINES AGENCY

6

## Scope

➢ **Use of the catalogue** to assess the suitability of data sources for specific studies

➢ **User guides**
  ➢ **Metadata definitions** and information expected to be used when describing a data source

  ➢ Dedicated sections for **inserting a new data source and maintenance** (to be further expanded once the system is live (e.g.: screenshots, technical guidance)

## Contents

7

EUROPEAN MEDICINES AGENCY

➢ **Two important aspects** of data quality of the primary data:

### Reliability

Based on e.g. the detection and correction of errors, missing data and implausible values (characteristic of the data source independent from its use for a specific study)

### Relevance

Data source to provide adequate and valid evidence informing a specific research question following the application of appropriate epidemiological and statistical techniques (may be required for some studies and not for others)
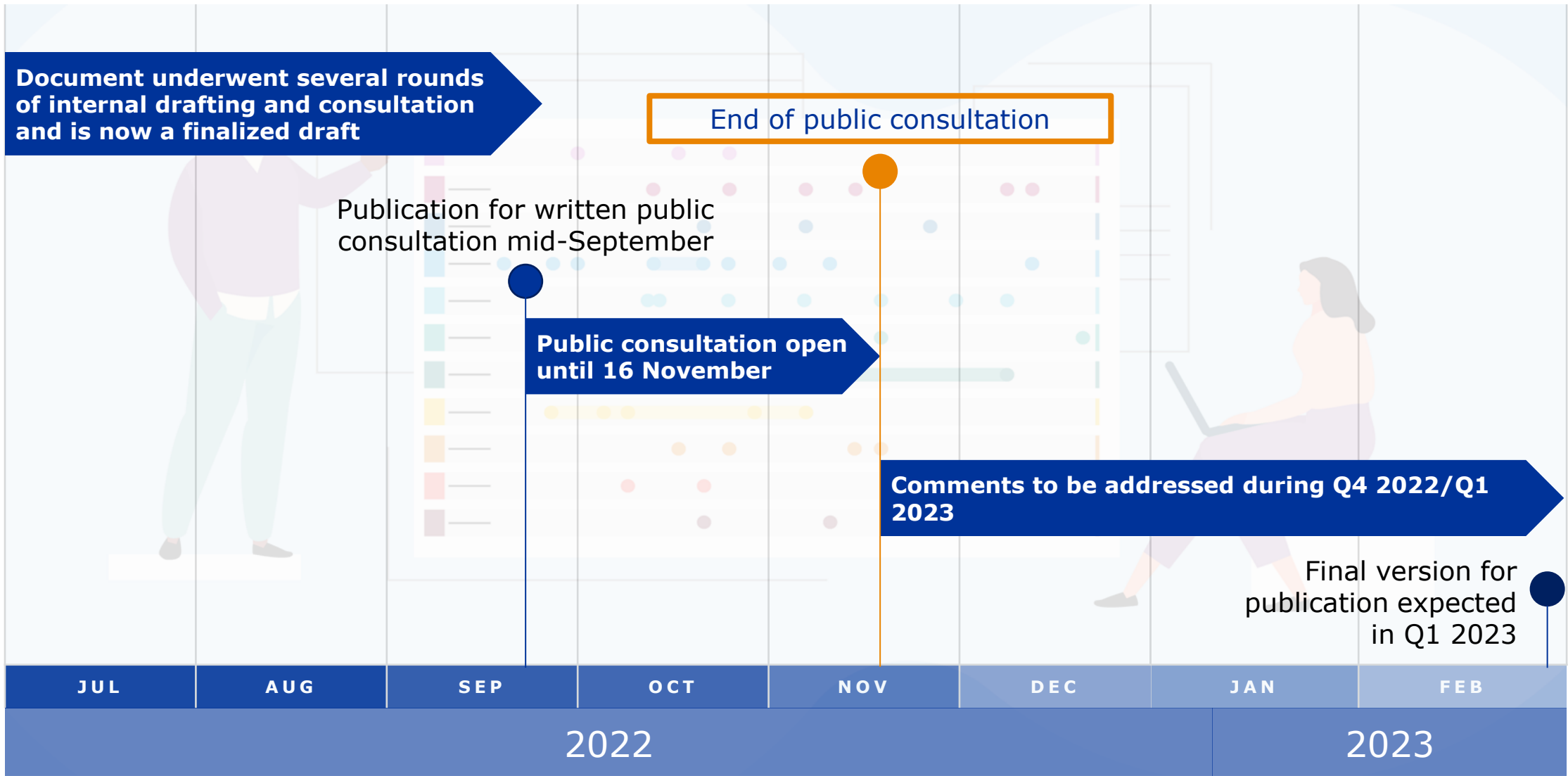
➢ The good practice guide describes the **specific variables collected** (metadata) to be used to specifically evaluate the data source's reliability and relevance

# Use cases described in the draft Good Practice Guide (examples)

| Use case | Content |
|---|---|
| *An investigator wants to identify suitable data sources for a planned study* | Outline of the steps recommended for a catalogue user in order to select a suitable data source for a particular study |
| *A study protocol submitted for a study mentions a data source and the assessor needs to understand the suitability of the data source proposed* | Description of the data elements useful for the assessment of the suitability of data source in the assessment of a study report |
| *An investigator writes a study reports for which they need to describe the data source(s) proposed to be used or used in the study* | Proposal of solutions when a comparison between the characteristics of several databases used in the study is difficult to perform |
| *A data holder wants to compare the features of a specific data source with other ones covering fully or partially the same population* | Recommendations on benchmarking several data sources – use of the harmonised description of the characteristics of each data source |

9

# Next steps – GPG drafting and publication

**Document underwent several rounds of internal drafting and consultation and is now a finalized draft**

End of public consultation

Publication for written public consultation mid-September

**Public consultation open until 16 November**

**Comments to be addressed during Q4 2022/Q1 2023**

Final version for publication expected in Q1 2023

| JUL | AUG | SEP | OCT | NOV | DEC | JAN | FEB |
|-----|-----|-----|-----|-----|-----|-----|-----|

| 2022 | 2023 |
|------|------|

# Some comments from the public consultation

**We received contributions from regulatory network, data holders, industry, medical associations, etc.**

**Thanks to all who contributed!**

## Comments were mainly related to:

### Catalogue maintenance and content validation
- Incentivisation of data holders to provide and update the metadata
- Quality check of data collected (how/when) to ensure data are entered in a harmonised way
- Management of the evolvement of metadata elements collected over time

### Scope of data source catalogue
- Clarify whether the catalogue will accommodate data sources outside the EU/EEA
- Clarify scope in terms of type of data sources (e.g., registries, electronic health records, claims data, m-health) and if there are exceptions

### Interoperability/collaboration with other catalogues (e.g., DARWIN EU®, EDHEN)
- Harmonisation of definitions and metadata terminology
- Minimisation of effort duplication for data holder

### Importance of data source accessibility
- In the process of identifying suitable data sources for a study this should be an early step
- Add in the catalogue information on accessibility, timelines and user restriction

### Interlink of Data quality framework (DQF) with the data sources catalogue and Good practice guide
- Alignment between Data quality framework and Good practice guide terminology and definitions
- Application of the Data Quality Framework to the catalogue

### Need for clarifications/definitions for specific metadata or proposals more metadata to collect
- Include more standardised vocabularies or use additional terminologies (e.g. to describe disease information)
- Definition needed  (e.g. for active population)
- Collection of more quantitative metadata / Producing quantitative descriptive statistics

EUROPEAN MEDICINES AGENCY

**Discovery and analysis phase** of the project to conclude by end of 2022

Testing to include **input from key stakeholders** (e.g., industry, ENCePP)

**Launch** is planned for Q4 2023 / Q1 2024

**Key features**:

| | |
|---|---|
| The **study catalogue and the data sources catalogue** will be linked → possibility to view information about data sources used in the study and vice versa | User **registration and login** will be required |
| Possibility to **search and export** information | Users will be able to **submit data through a webform** (insert/update content) |
| Studies will be visible in the relevant **medicines overview page** in the EMA website | The approval/rejection of data will trigger **e-mail notifications to user submitting data** (i.e.: the EMA administration functions) |
| **Globally unique and persistent identifiers** according to FAIR (Findable, Accessible, Interoperable, and Re-usable) data principles | **Notifications will be sent prompting users** to update their records |

- With regards to **current ENCePP:**

  - Data from **ENCePP Resources database and EU PAS Register will be migrated**, cleaned-up, transformed, enhanced and published in the new catalogue

  - **EMA will reach out in Q1 2023 to data holders** of data sources included in ENCePP Resources database in order to ask the update and completion of the (new) metadata

  - Both **data source fields and studies are changing significantly** (data sources – much more information; studies – some fields are added, many are transformed)

  - All the **functionality existing in ENCePP** will be implemented and enhanced in the new system

13

# Any questions?

Email: [metadata@ema.europa.eu](mailto:metadata@ema.europa.eu)

**Official address**  Domenico Scarlattilaan 6  •  1083 HS Amsterdam  •  The Netherlands
**Telephone** +31 (0)88 781 6000
**Send us a question**  Go to www.ema.europa.eu/contact

Follow us on 🐦 **@EMA_News**