EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

# From Data to Decision

The European Medicines Regulatory Network initiatives to leverage data to support decision-making

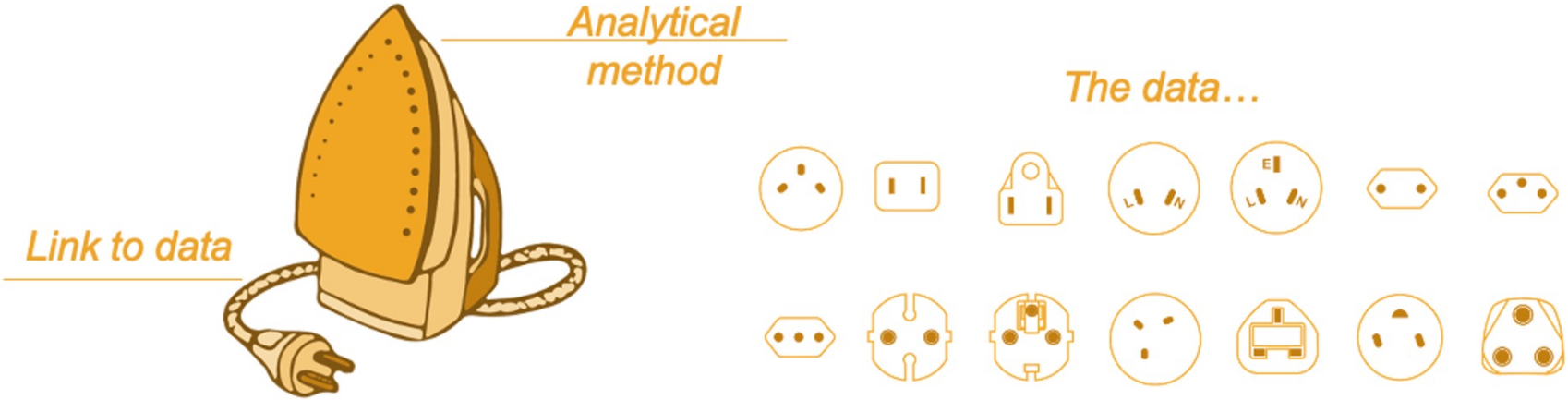
EJPRD Industry Webinar

Presented by Luis Pinheiro
Slides from Luis Pinheiro, Catherine Cohet, Kelly Plueschke, Andrej Segec, and others in TDA Data Analytics and Methods Taskforce

# Content

- Standardisation of data to unlock greater value from real-world evidence
- Real-world evidence in regulatory decision-making
- Generation of real-world evidence
- Data discoverability and quality
- European Medicines Regulatory Network AI drivers & approach

# Improving interoperability of data

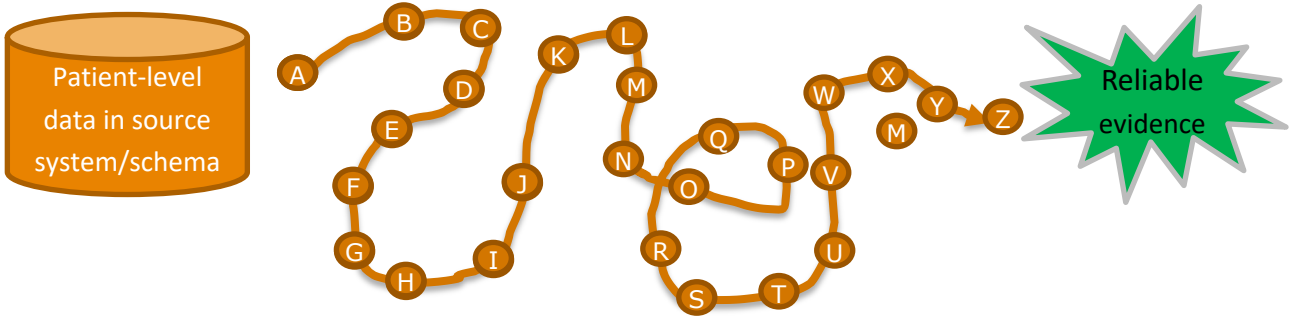



- Increasing productivity to an industrial level requires the automation of the analytical processes, which in turn cannot be done without a rigorous standard representation of the data.

- Full interoperability of the data is needed with respect to structure (syntactic interoperability) and coding systems (semantic interoperability) by using a Common Data Model (CDM)

# Generating Reliable Evidence using a Common Data Model

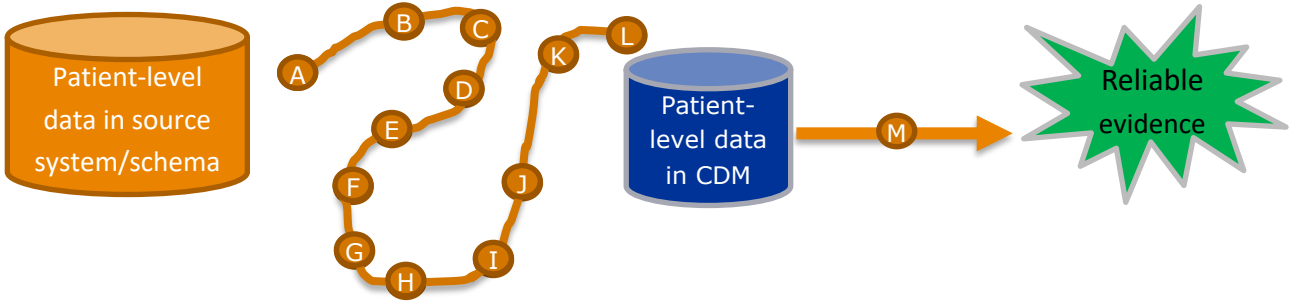
We need to make studies repeatable, reproducible, replicable, generalisable, and robust




A Common Data Model will enable standardised analytics to generate reliable evidence.

# Generating Reliable Evidence using a Common Data Model

We need to make studies repeatable, reproducible, replicable, generalisable, and robust




A Common Data Model will enable standardised analytics to generate reliable evidence.

# Use or RWE in marketing authorisation applications

EUROPEAN MEDICINES AGENCY

| | Flynn et al. (2021) What was the Contribution of Real-World Evidence in EU? | Eskola et. al (2021) Use of Real-World Data and Evidence in Drug Development in EU | Purpura et al. (2021) The Role of Real-World Evidence in FDA |
|---|---|---|---|
| No. of products reviewed | 158 | 111 | 136 |
| Period | Jan 2018 – Dec 2019 (submitted MAAs, incl. non-published info) | Jan 2018 – Dec 2019 (approved MAAs, only published info) | Jan 2019 – June 2021 (approved MAAs, only published info) |
| No. of products with RWE included in MAA | **63 (39.9%)** | **111 (100%)** | **116 (85.2%)** |
| Therapeutic area | Oncology and anti-infectives | Oncology, haematology and anti-infectives | Oncology and anti-infectives |
| Key messages | • Widespread use of RWE<br>• RWE in pre-authorisation (1/3), post-authorisation (2/3)<br>• RWE to support safety (87%) and efficacy (49%)<br>• Most common data sources: registries (60%), hospital data (32%) | RWE contribution to learnings and regulatory decisions:<br>• in virtually all phases<br>• across ≠ therapeutic areas<br>• with ≠ product characteristics<br>RWE particularly supports conditional MA and approval of orphan medicines | Successful RWE requires:<br>• fit-for-purpose data<br>• robust study design, appropriate data collection, thoughtful data analysis<br>• proactive communication (with FDA) |

# How can RWE support regulatory decision-making?

**Understand the clinical context**

- Disease epidemiology
- Clinical management
- Drug utilisation

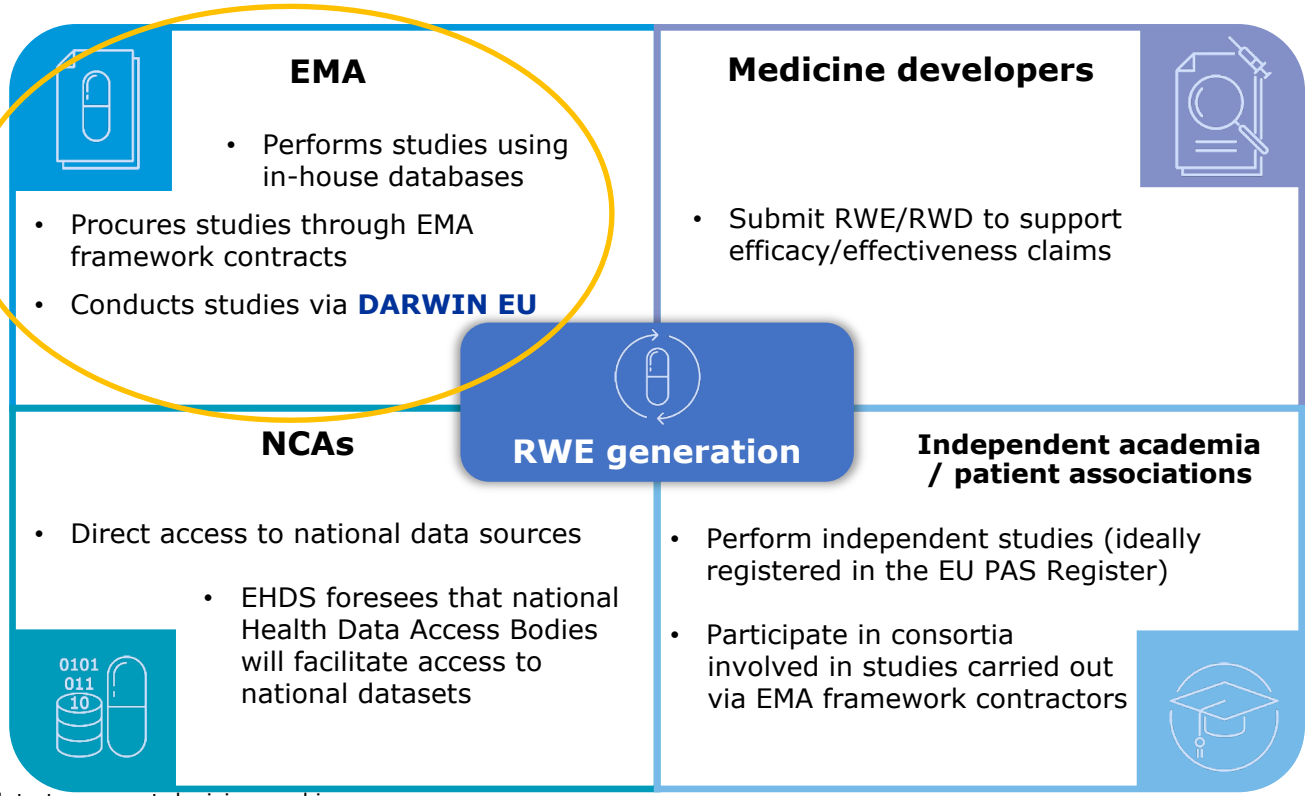
**Support the planning and validity**

- Design and feasibility of planned studies
- Representativeness and validity of completed studies

**Investigate associations and impact**

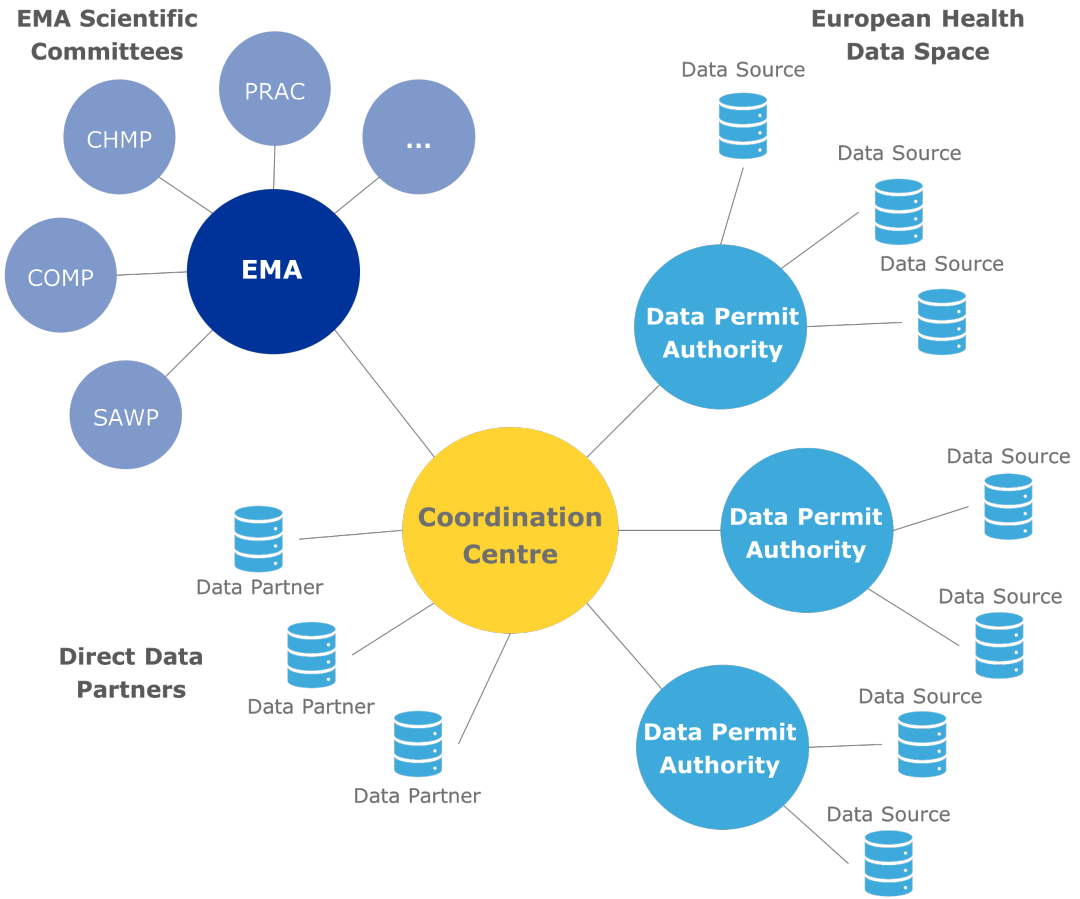
- Effectiveness and safety studies
- Impact of regulatory actions

# Who generates RWE for EU regulatory purposes?




Leveraging data to support decision-making

DARWIN EU® is a federated **network** of **data**, **expertise** and **services** that supports better decision-making throughout the product lifecycle by generating reliable **evidence from real world healthcare data**




## FEDERATED NETWORK PRINCIPLES

- Data stays **local**
- Use of OMOP **Common Data Model** (where applicable) to perform studies in a timely manner and increase consistency of results

Leveraging data to support decision-making

# Who will benefit from DARWIN EU ?

EUROPEAN MEDICINES AGENCY

## EU medicines regulators

**Understand the clinical context**

**Support the planning and validity**

**Investigate associations and impact**

DARWIN EU® will **increase the capacity** of the EMRN to undertake high-quality observational studies based on RWD and **reduce the time** per study

**EU patients and healthcare professionals**
Faster access to innovative medicines and safe and effective use

**European Commission**
Key use case for the European Health Data Space

**National competent authorities**
Support health policy and delivery of healthcare systems

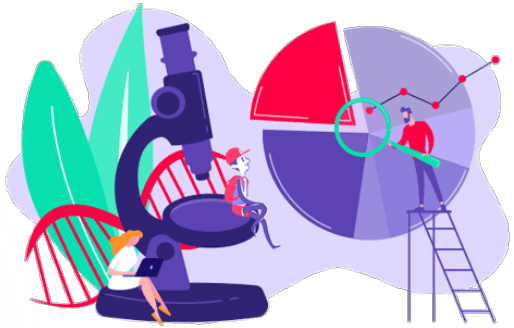
**HTA bodies and payers**
Support better quality decisions including on cost-effectiveness

**EU and international health agencies**
Use cases specific for other EU Agencies such as ECDC

**Academia and research organisations**
Increase use of RWE, methodology development, and better data quality

**Industry**
Enable better evidence supporting decision-making, increase receptiveness for RWE in MAAs, and reduce time & cost of medicines development

# Rare diseases context



frontiers | Frontiers in Pharmacology

TYPE Review
PUBLISHED 04 August 2022
DOI 10.3389/fphar.2022.924648

Check for updates

OPEN ACCESS

EDITED BY
Virginie Hivert,
European Organisation for Rare Diseases (EURORDIS), France

REVIEWED BY
Veronica Miller,
University of California, Berkeley, United States
Chiuhui Mary Wang,
Rare Diseases International, France
Sabina Antonela Antoniu,
Grigore T. Popa University of Medicine and Pharmacy, Romania
Marc Marie Dooms,
University Hospitals Leuven, Belgium

## Contribution of patient registries to regulatory decision making on rare diseases medicinal products in Europe

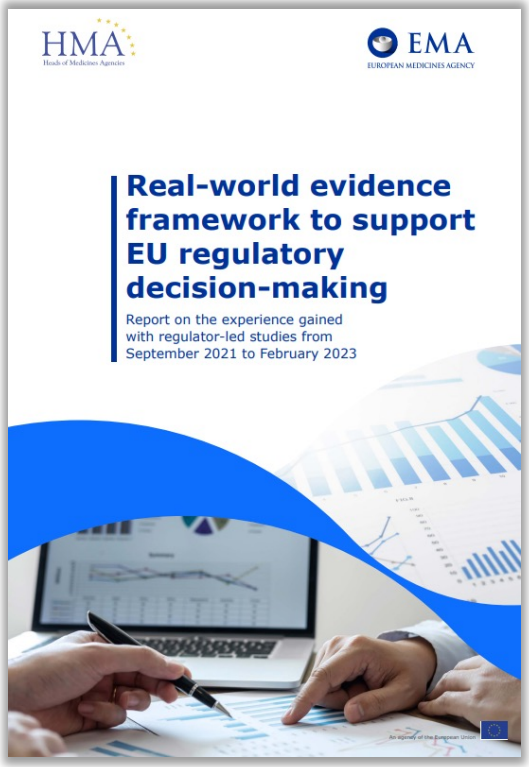
Carla J. Jonker[1,2], Elisabeth Bakker[1,3], Xavier Kurz[1] and Kelly Plueschke[1]*

[1]European Medicines Agency (EMA), Amsterdam, Netherlands, [2]Dutch Medicines Evalua (CBG-MEB), Utrecht, Netherlands, [3]Department of Clinical Pharmacy and Pharmacolog Medical Center Groningen (UMCG), Groningen, Netherlands

- Integrating registry data in medicines' life cycle
- Increasing use of RWE for regulatory purposes
- Use cases demonstrating benefits of registries for regulatory purposes
- Challenges
- European initiatives promoting registries for regulatory purposes

*Jonker et al., Front. Pharmacol., 04 August 2022; Sec. Drugs Outcomes Research and Policies*
https://doi.org/10.3389/fphar.2022.924648

# Review of EMA-conducted RWE studies




Leveraging data to support decision-making

- Objective: to take stock of the experience with regulatory-led RWE studies and evaluate opportunities and challenges in supporting regulatory decision-making
- Includes a portfolio of use cases

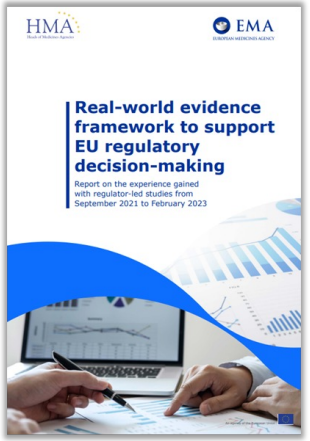
→ **RWE studies able to address a broad range of research questions**

→ Feasibility remains a frequent challenge, esp. if conditions/medicines not in primary care, medicines not prescribed in country of study, rare diseases, if tight procedural timelines

LINK to report

# Review of EMA-conducted RWE studies

## Rare disease recommendations




- There is a need to **expand the access to diverse and complementary data sources** (e.g., hospital health records, registries) that are:
  a) **suitable for addressing research questions on rare diseases, or conditions treated, and medicines prescribed at secondary care level,** and
  b) with enough granularity to have access to the relevant information to fill the current evidence gaps.

Recommendation: More than half (8/14) of the research requests were related to rare diseases or specialist settings or would have required richer information as the one available, e.g., in disease-specific **registries or biobanks**. This should be taken into account with a view to contracting future in-house data sources and onboarding of new DARWIN EU® data partners.

Recommendations: Use of a **curated and validated library of phenotypes** to identify rare conditions will help to address the risk of misclassification. In addition, **clinical experts** should be involved in the definition and validation. Notably, DARWIN EU® is building a phenotype library that will be curated by medical experts and periodically validated, which has the potential to improve rare diseases identification.

# There are many challenges… that can be overcome

EUROPEAN MEDICINES AGENCY

## Operational:

linked to feasibility, governance and sustainability issues

## Technological:

different terminologies, case definitions, data formats, coding systems, quality and content

## Methodological:

original purpose of data collection may not be for research

*Non exhaustive list*

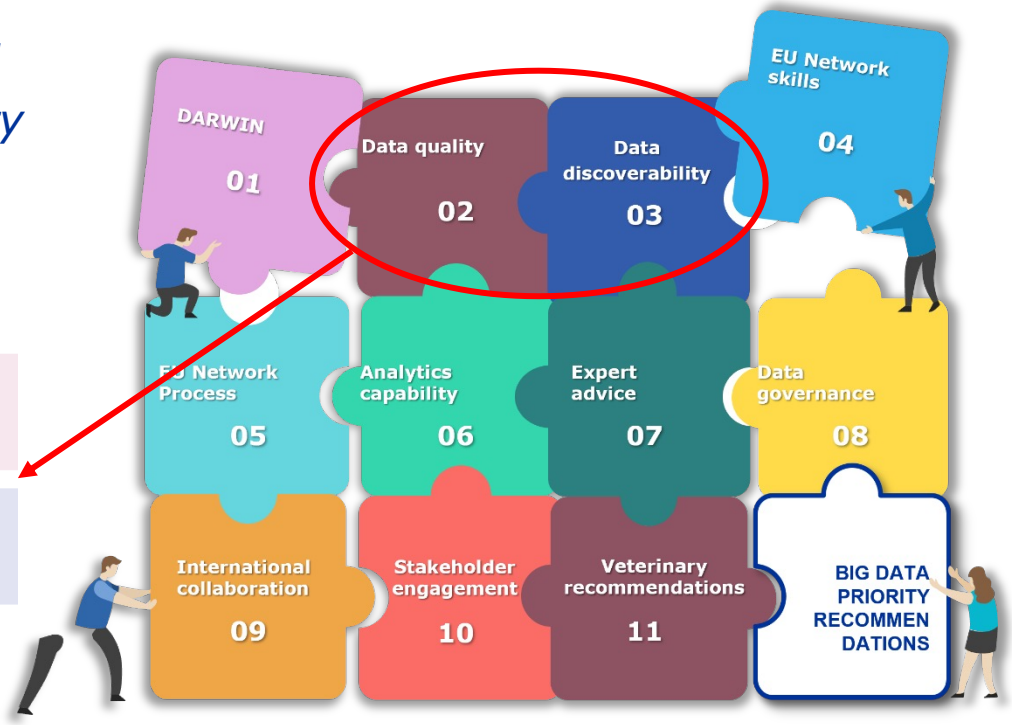
**Solutions**

- **Open dialogue** → Share needs on data and manage expectations (e.g., PRIME, Qualification procedure)
- **Publication of better description** of RWD sources, processes and policy for collaboration and data sharing → Discoverability of "Fit-for-Purpose" data sources
- Multi-stakeholder **agreement on common sets of data** elements/terminologies/format (e.g., ERICA, EJP RD)
- Enabling **common data models, interoperability** (EHDEN / OMOP / DARWIN / Horizon Europe projects)
- **European Health Data Space** (EHDS)
- **Compliance** with the best methodological standards (e.g., EMA Guideline on Registry-based studies, ENCePP guide, Data Quality Framework)
- **Continuous engagement** with HCP/patients for "return on investment"; appropriate trainings

# HMA/EMA Big Data Steering Group workplan 2022-2025

*Framework to enable use of data and facilitate its integration into regulatory decision-making*

| | |
|---|---|
| **DATA QUALITY AND REPRESENTATIVINESS** | **Data quality framework** roll-out<br>**Good practices** on regulatory data science, management and software<br>Strengthen **data qualification**<br>Activity links to **EHDS** |
| **DATA DISCOVERABILITY** | Publish **real-world data and studies catalogues**<br>Explore **patient experience** data analysis<br>**Clinical trials protocol** analytics<br>Review utility of **eHealth data** and **social media** |




Leveraging data to support decision-making

# Data quality framework

HMA
Heads of Medicines Agencies

EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

30 October 2023
Data Analytics and Methods Task Force
EMA/326985/2023

## Data Quality Framework for EU medicines regulation

| | |
|---|---|
| **Draft agreed by BDSG for release for consultation** | 10 October 2022 |
| **End of consultation (deadline for comments)** | 18 November 2022 |
| **Agreed by BDSG and MWP** | 30 June 2023 |
| **Adopted by CHMP** | 30 October 2023 |

| | |
|---|---|
| **Keywords** | Data quality framework, medicines regulation, data quality dimensions, primary and secondary use of data |

# Artificial intelligence | Drivers

## Process Analytics

Internal

Improve **efficiency** by automating and digitalizing processes

## Regulatory submissions

External

**Regulate applications of AI** in medicines with a view to help create value for public health

## Healthcare data analytics

Internal

**Structure information and increase insights** into data to inform decision-making

## External collaboration

Be an effective collaborator, incl. on legislation development

- FDA, ICMRA, WHO, CIOMS, EU Agencies, etc.

Leveraging data to support decision-making

# European Medicines Regulatory Network | Key initiatives

**Key initiatives of the European Medicines Regulatory Network**

Big Data Task Force | European Medicines Regulatory Network Strategy | Regulatory Science Strategy

**HMA/EMA AI Workshop 2021**

## Guidance

Offer clear, balanced, guidance to developers, marketing authorisation holders, etc.

## Upskilling

Provide training across the network to foster competent and responsible use/regulation of AI

## Collaborations

Leverage collaborations to improve knowledge, reduce uncertainty and facilitate alignment

# AI Reflection Paper


13 July 2023
EMA/CHMP/CVMP/83833/2023
Committee for Medicinal Products for Human Use (CHMP)
Committee for Medicinal Products for Veterinary Use (CVMP)

## Reflection paper on the use of Artificial Intelligence (AI) in the medicinal product lifecycle

Draft

| | |
|---|---|
| Draft agreed by Committee for Medicinal Products for Human Use (CHMP) Methodology Working Party | July 2023 |
| Draft adopted by CVMP for release for consultation | 13 July 2023 |
| Draft adopted by CHMP for release for consultation | 10 July 2023 |
| Start of public consultation | 19 July 2023 |
| End of consultation (*deadline for comments*) | 31 December 2023 |

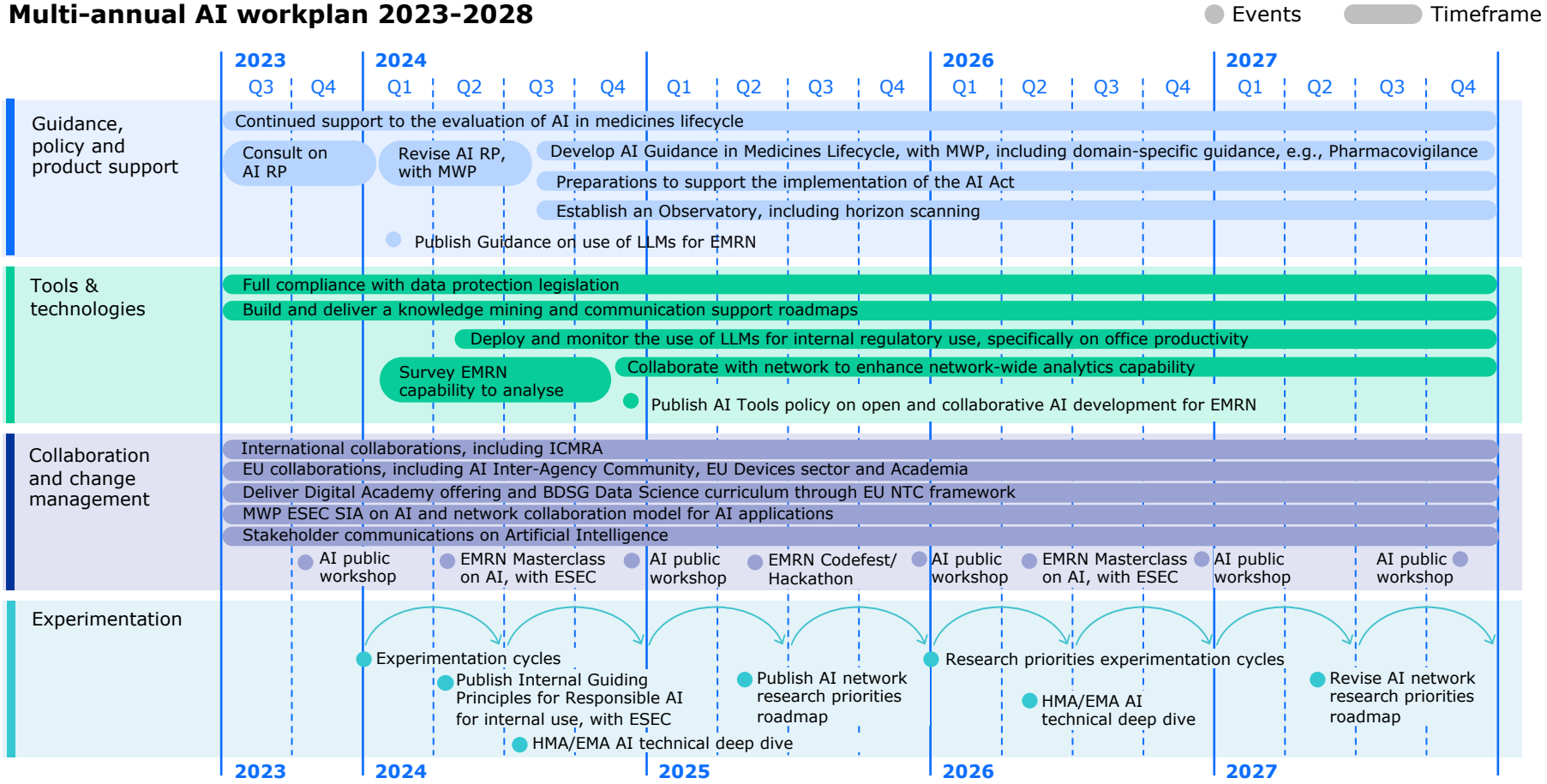
Comments should be provided using this EUSurvey form. For any technical issues, please contact the EUSurvey Support.

| | |
|---|---|
| Keywords | *Artificial intelligence, AI, machine learning, ML, regulatory, medicine, human medicinal product, veterinary medicinal product* |

**Public consultation to 31 Dec 2023**
**AI Workshop 20 & 21 Nov 2023**

Leveraging data to support decision-making

# Multi-annual AI workplan 2023-2028



# Experimentation | Analytic vs Generative AI

| | Analytic AI (Aim: group similar things or be accurate) | Generative AI (Aim: mimic) |
|---|---|---|
| Capabilities | Analyse, recognise patterns, predict, probabilistic classification, etc. | Generate content, structure text, support cognitive work. *Derives from Analytic AI.* |
| General use | Automation, analysis, routine decision making, domain-specific interaction | Support cognitive work, content generation, querying and structuring data |
| Input | Usually structured data | Usually unstructured data |

Leveraging data to support decision-making

# Data insights | Clinical prediction models

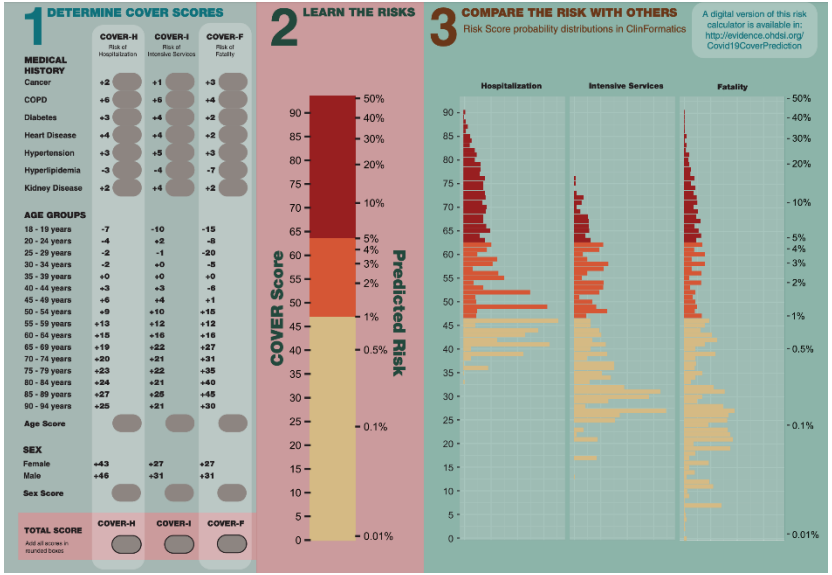
## Description

Algorithmic risk scores not new to medicine and public health. ML models add the possibility to **predict probability of a clinical outcome** purely from a data-driven perspective

## Pharmacoepidemiology use cases

Prediction of the probability of a clinical outcomes or behaviour (e.g. risk of abuse of opioids)

## Example




Seek COVER: Development and validation of a personalised risk calculator for COVID-19 outcomes in an international network

# Data insights | Probabilistic phenotyping

## Description

In probabilistic phenotyping the ML identifies patterns in the characteristics of a population of interest and **predicts the likelihood that a patient is an element of that population**

## Pharmacoepidemiology use cases

Cohort development – populations that share similar traits, particularly for phenotypes that have complex logic or are not directly identifiable from existing diagnostic/laboratory codes (e.g., gestational age)

## Example

An Application of Machine Learning in Pharmacovigilance: Estimating Likely Patient Genotype From Phenotypical Manifestations of Fluoropyrimidine Toxicity

Luis Correia Pinheiro[1,*], Julie Durand[1] and Jean-Michel Dogné[2,3]

Dihydropyrimidine dehydrogenase (DPD)-deficient patients might only become aware of their genotype after exposure to dihydropyrimidines, if testing is performed. Case reports to pharmacovigilance databases might only contain phenotypical manifestations of DPD, without information on the genotype. This poses a difficulty in estimating the cases due to DPD. Auto machine learning models were developed to train patterns of phenotypical manifestations of toxicity, which were then used as a surrogate to estimate the number of cases of DPD-related toxicity. Results indicate that between 8,878 (7.0%) and 16,549 (13.1%) patients have a profile similar to DPD deficient status. Results of the analysis of variable importance match the known end-organ damage of DPD-related toxicity, however, accuracies in the range of 90% suggest presence of overfitting, thus, results need to be interpreted carefully. This study shows the potential for use of machine learning in the regulatory context but additional studies are required to better understand regulatory applicability.

# Exposing data | Generative AI (GenAI)

## Description

**Large language models,** facilitate extraction and structuring of information, particularly that which is extensively available in the training data.

**LLMs** can be used for NER have limitations

## Pharmacoepidemiology use cases

Structuring data to be ingested, e.g. in a protocol.

Coding

Screening through information

## Example

json Copy code

```json
{
  "description": "Anaphylaxis is a severe allergic reaction that can be
  "symptoms": {
    "skin": ["hives", "itching", "redness"],
    "respiratory": ["shortness of breath", "wheezing", "coughing"],
    "gastrointestinal": ["nausea", "vomiting", "diarrhea"],
    "cardiovascular": ["weak pulse", "palpitations", "dizziness"],
    "other": ["anxiety", "confusion", "loss of consciousness"]
  }
}
```

sql Copy code

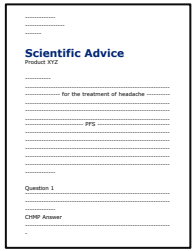
```sql
SELECT sex, COUNT(pat_id) AS patient_count
FROM person
GROUP BY sex;
```

# Scientific Explorer / From multiple unstructured sources to a single structured target

EUROPEAN MEDICINES AGENCY

## Data Sources

## Ingestion and Pre-processing




Scraping Websites
Ingest .doc .docx .pdf…

## Extracted Terms

```
{
   Doc Type :
Scientific Advice,
   Product : XYZ,
   Indication :
Headache,
   Primary Endpoint
: PFS,
   …
}
```

Pattern Matching
NER models

## Search engine & UI

**Lakehouse**

Index, store and share structured data

## Large Language Models

Engineered Prompt

```
Extract the following set of information from the
letter. For each extraction …

- Comparator: The Comparator, control, reference group,
cohort or arm. A few relevant example are placebo,
active comparator, low dose, historical, standard of
care, sham, vehicle, uncontrolled, noncomparative,
intrapatient, within patient
- Primary Endpoint(s): …
```

GPT-4 Output

```
Trial : Paediatric IL-23 platform trial
- Comparator
  = Proposal by the Applicant : Using adult placebo as
an external control group
  = CHMP level of Agreement : Partially agreed
  = CHMP Caveats : Comparison should be adequately
adjusted for all possible relevant differences between
patient populations.
  = References : Question 2
```

**$F_1$ = .90**
Precision = .92
Recall = .89

From
~ 300 validations

Leveraging data to support decision-making

# Any questions?

## Further information

luis.pinheiro@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**