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

Leveraging data to support decision-making
• 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)

Leveraging data to support decision-making
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.

Leveraging data to support decision-making.
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.

Leveraging data to support decision-making
Use or RWE in marketing authorisation applications

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of products reviewed</td>
<td>158</td>
<td>111</td>
<td>136</td>
</tr>
<tr>
<td>Period</td>
<td>Jan 2018 – Dec 2019 (submitted MAAs, incl. non-published info)</td>
<td>Jan 2018 – Dec 2019 (approved MAAs, only published info)</td>
<td>Jan 2019 – June 2021 (approved MAAs, only published info)</td>
</tr>
<tr>
<td>No. of products with RWE included in MAA</td>
<td><strong>63 (39.9%)</strong></td>
<td><strong>111 (100%)</strong></td>
<td><strong>116 (85.2%)</strong></td>
</tr>
<tr>
<td>Therapeutic area</td>
<td>Oncology and anti-infectives</td>
<td>Oncology, haematology and anti-infectives</td>
<td>Oncology and anti-infectives</td>
</tr>
<tr>
<td>Key messages</td>
<td>• Widespread use of RWE</td>
<td>RWE contribution to learnings and regulatory decisions:</td>
<td>Successful RWE requires:</td>
</tr>
<tr>
<td></td>
<td>• RWE in pre-authorisation (1/3), post-authorisation (2/3)</td>
<td>• in virtually all phases</td>
<td>• fit-for-purpose data</td>
</tr>
<tr>
<td></td>
<td>• RWE to support safety (87%) and efficacy (49%)</td>
<td>• across ≠ therapeutic areas</td>
<td>• robust study design, appropriate data collection, thoughtful data analysis</td>
</tr>
<tr>
<td></td>
<td>• Most common data sources: registries (60%), hospital data (32%)</td>
<td>• with ≠ product characteristics</td>
<td>• proactive communication (with FDA)</td>
</tr>
<tr>
<td></td>
<td>Leverage data to support decision-making</td>
<td>RWE particularly supports conditional MA and approval of orphan medicines</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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

Leveraging data to support decision-making
Who generates RWE for EU regulatory purposes?

**EMA**
- Performs studies using in-house databases
- Procures studies through EMA framework contracts
- Conducts studies via DARWIN EU

**Medicine developers**
- Submit RWE/RWD to support efficacy/effectiveness claims

**NCAs**
- Direct access to national data sources
  - EHDS foresees that national Health Data Access Bodies will facilitate access to national datasets

**Independent academia / patient associations**
- Perform independent studies (ideally registered in the EU PAS Register)
- Participate in consortia involved in studies carried out via EMA framework contractors

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?

**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

Leveraging data to support decision-making

**Classified as public by the European Medicines Agency**
Rare diseases context

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

Carla J. Jonker, Elisabeth Bakker, Xavier Kurz and Kelly Plueschke

European Medicines Agency (EMA), Amsterdam, Netherlands; Dutch Medicines Evaluation Board (CBG-MEB), Utrecht, Netherlands; Department of Clinical Pharmacy and Pharmacology, 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

Leveraging data to support decision-making

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

Classified or public by the European Medicines Agency
Review of EMA-conducted RWE studies

- 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

**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

Leveraging data to support decision-making
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 REPRESENTATIVENESS**
- **Data quality framework** roll-out
  - Good practices on regulatory data science, management and software
  - Strengthen data qualification
  - Activity links to EHDSS

**DATA DISCOVERABILITY**
- **Publish real-world data and studies catalogues**
  - Explore patient experience data analysis
  - Clinical trials protocol analytics
  - Review utility of eHealth data and social media

Leveraging data to support decision-making
Data quality framework

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**
  - Improve **efficiency** by automating and digitalizing processes

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

- **Healthcare data analytics**
  - 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

Leveraging data to support decision-making
AI Reflection Paper

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

<table>
<thead>
<tr>
<th>Event</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Draft agreed by Committee for Medicinal Products for Human Use (CHMP) Methodology Working Party</td>
<td>July 2023</td>
</tr>
<tr>
<td>Draft adopted by CVMP for release for consultation</td>
<td>13 July 2023</td>
</tr>
<tr>
<td>Draft adopted by CHMP for release for consultation</td>
<td>10 July 2023</td>
</tr>
<tr>
<td>Start of public consultation</td>
<td>19 July 2023</td>
</tr>
<tr>
<td>End of consultation (deadline for comments)</td>
<td>31 December 2023</td>
</tr>
</tbody>
</table>

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

Leveraging data to support decision-making

Public consultation to 31 Dec 2023
AI Workshop 20 & 21 Nov 2023
## Multi-annual AI workplan 2023-2028

### Guidance, policy and product support
- **2023 Q3-Q4**: Continued support to the evaluation of AI in medicines lifecycle
  - Consult on AI RP
- **2024 Q1-Q4**: Revise AI RP, with MWP
  - Develop AI Guidance in Medicines Lifecycle, with MWP, including domain-specific guidance, e.g., Pharmacovigilance
- **2026 Q1-Q4**: Preparations to support the implementation of the AI Act
  - Establish an Observatory, including horizon scanning
- **2027 Q1-Q4**: Publish Guidance on use of LLMs for EMRN

### Tools & technologies
- **2023 Q3-2024 Q4**: Full compliance with data protection legislation
  - Build and deliver a knowledge mining and communication support roadmaps
- **2024 Q1-2027 Q4**: Deploy and monitor the use of LLMs for internal regulatory use, specifically on office productivity
  - Survey EMRN capability to analyse
  - Collaborate with network to enhance network-wide analytics capability
- **2023 Q3-2025 Q4**: Publish AI Tools policy on open and collaborative AI development for EMRN

### Collaboration and change management
- **2023 Q3-2024 Q4**: International collaborations, including ICMRA
  - EU collaborations, including AI Inter-Agency Community, EU Devices sector and Academia
- **2024 Q1-2027 Q4**: Deliver Digital Academy offering and BDSG Data Science curriculum through EU NTC framework
  - MWP ESEC SIA on AI and network collaboration model for AI applications
- **2023 Q3-2025 Q4**: Stakeholder communications on Artificial Intelligence

### Experimentation
- **2023-2027**: Experimentation cycles
  - Publish Internal Guiding Principles for Responsible AI for internal use, with ESEC
  - HMA/EMA AI technical deep dive
  - Publish AI network research priorities roadmap
  - Research priorities experimentation cycles
  - HMA/EMA AI technical deep dive
  - Revise AI network research priorities roadmap
Experimentation | Analytic vs Generative AI

**Analytic AI**
(Aim: group similar things or be accurate)
- Analyse, recognise patterns, predict, probabilistic classification, etc.
- Automation, analysis, routine decision making, domain-specific interaction
- Usually structured data

**Capabilities**

**Generative AI**
(Aim: mimic)
- Generate content, structure text, support cognitive work. Derives from Analytic AI.
- Support cognitive work, content generation, querying and structuring data
- Usually unstructured data

**General use**

**Input**

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 Pinheiro¹, Julie Durand¹ and Jean-Michel Dogne²,³

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.

Leveraging data to support decision-making
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

Leveraging data to support decision-making

**Example**

```
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"],
}
```
**Scientific Explorer** / From multiple unstructured sources to a single structured target

**Data Sources**

- Scraping Websites
- Ingest .doc .docx .pdf...

**Ingestion and Pre-processing**

**Extracted Terms**

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

**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

Leveraging data to support decision-making

**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

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