Big Data Workplan 2023-2025

HMA/EMA joint Big Data Steering Group

Contents

Introduction

The vision on Big Data is a strengthened regulatory system that can efficiently integrate data analysis into its assessment processes to improve decision making. Knowing when and how to have confidence in novel technologies and the evidence generated from Big Data will benefit public health by accelerating medicines development, improving treatment outcomes and facilitating earlier patient access to new treatments.

The 4th joint HMA-EMA Big Data Steering Group (BDSG) workplan was adopted in June 2023 and covers activities until 2025.

This document introduces each topic and outlines key deliverables. The plan was informed by stakeholder and expert consultation. The document is structured in line with the key recommendations of the Big Data Task Force (see Annex I). The scope of activities under these recommendations covers mainly human medicines, however veterinary aspects are included when relevant.
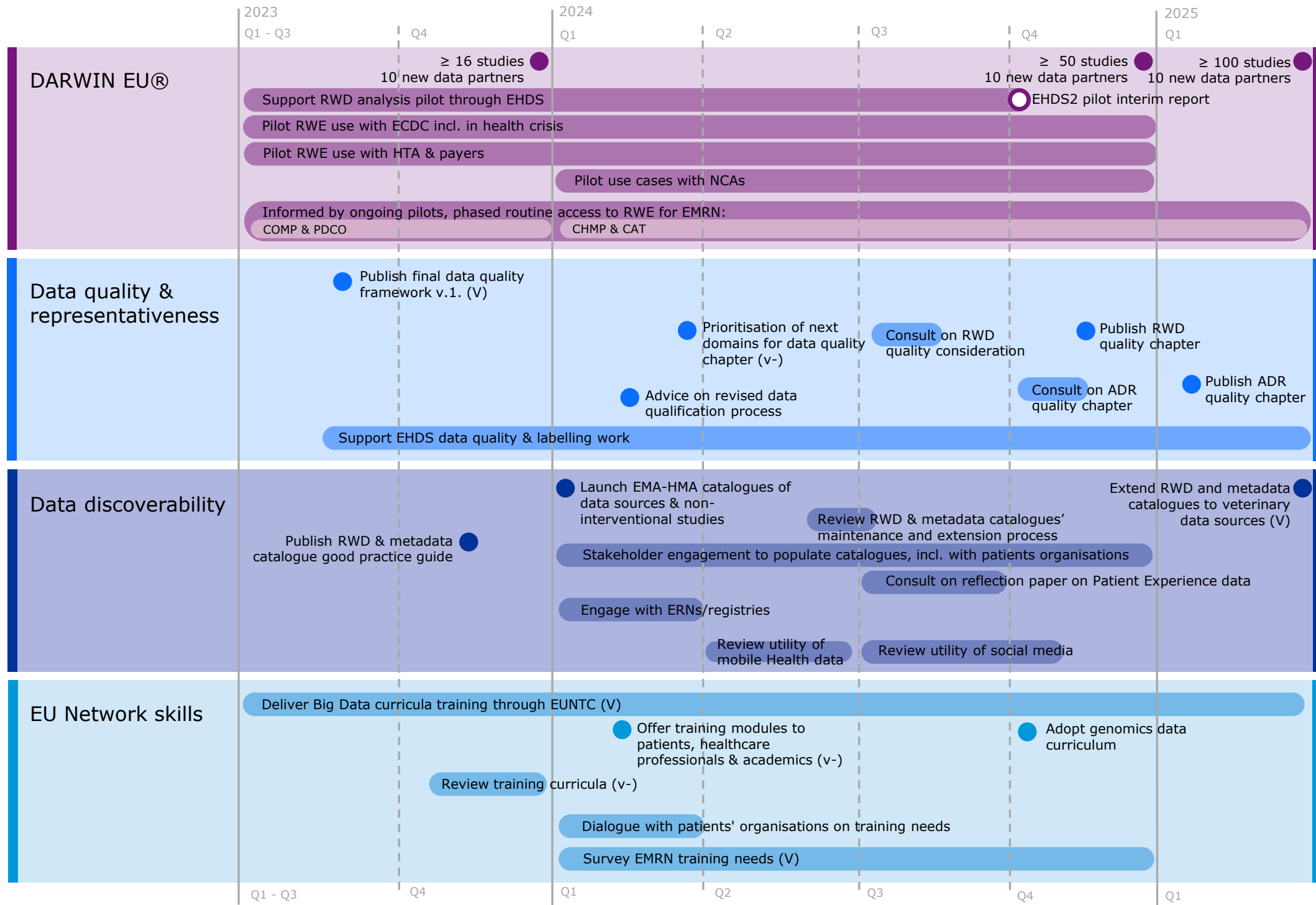
Progress is delivered in line with the agreed workplan and the Network Strategy to 2025 to enable the use and establish the value of the Big Data in the development, authorisation and supervision of medicines in EU.
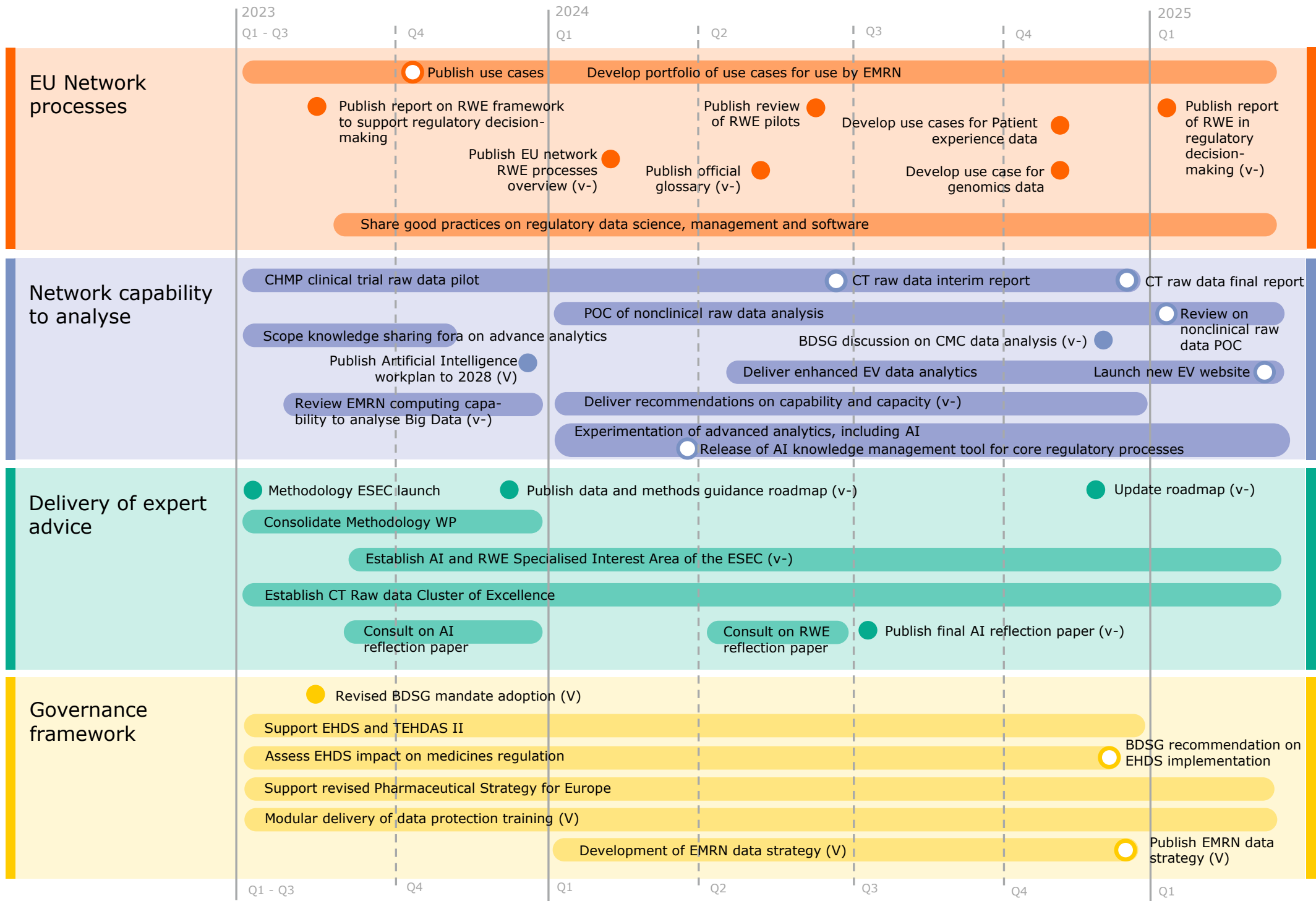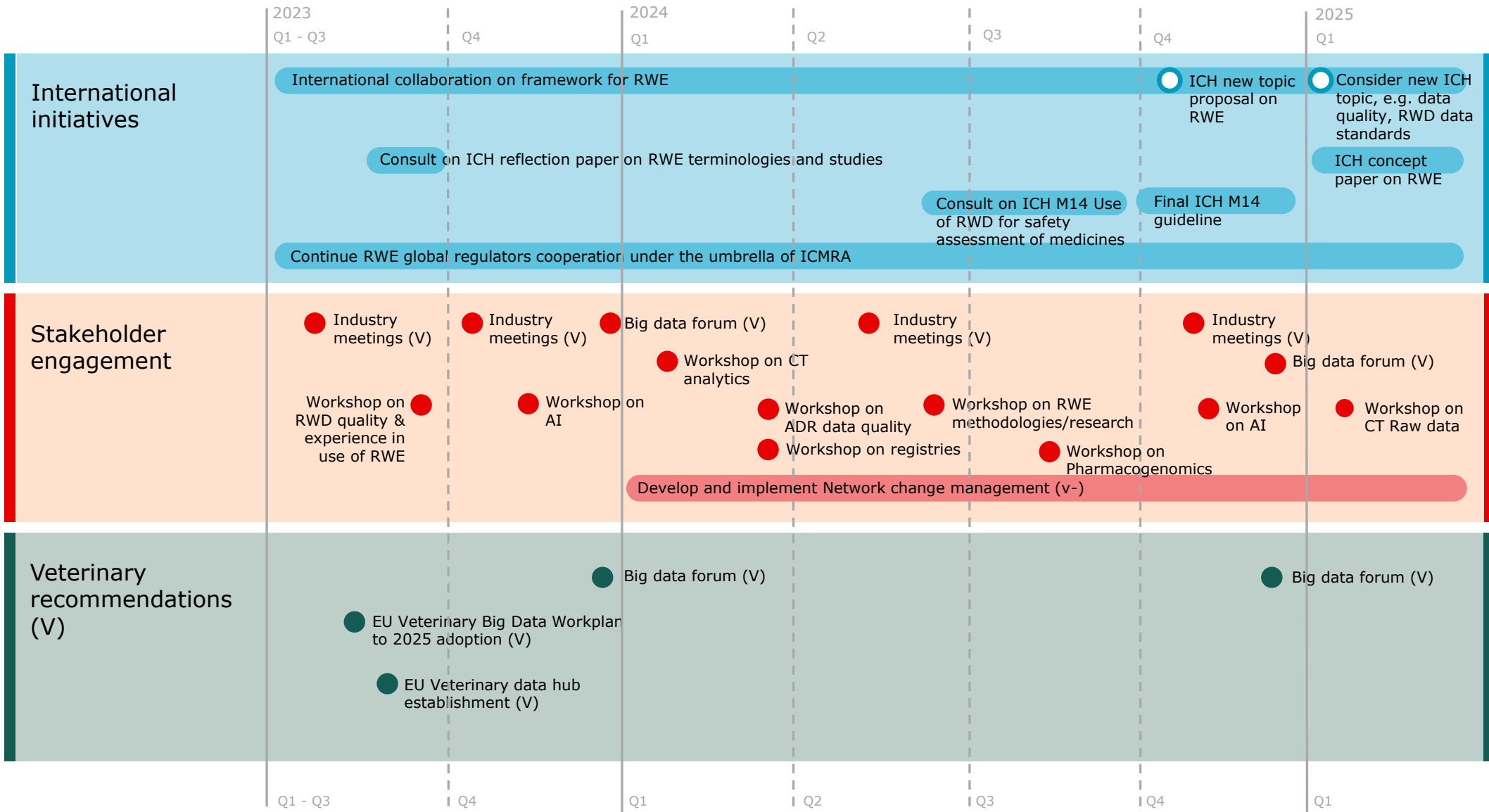
Timeline overview

BDSG workplan 2023-2025

*Scope applicable to veterinary aspects ((v-) if tbc)

| | 2023 Q1 - Q3 | Q4 | 2024 Q1 | Q2 | Q3 | Q4 | 2025 Q1 |

DARWIN EU®

≥ 16 studies
10 new data partners

≥ 50 studies
10 new data partners

≥ 100 studies
10 new data partners

Support RWD analysis pilot through EHDS

EHDS2 pilot interim report

Pilot RWE use with ECDC incl. in health crisis

Pilot RWE use with HTA & payers

Pilot use cases with NCAs

Informed by ongoing pilots, phased routine access to RWE for EMRN:
COMP & PDCO   CHMP & CAT

Data quality & representativeness

Publish final data quality framework v.1. (V)

Prioritisation of next domains for data quality chapter (v-)

Consult on RWD quality consideration

Publish RWD quality chapter

Advice on revised data qualification process

Consult on ADR quality chapter

Publish ADR quality chapter

Support EHDS data quality & labelling work

Data discoverability

Launch EMA-HMA catalogues of data sources & non-interventional studies

Extend RWD and metadata catalogues to veterinary data sources (V)

Review RWD & metadata catalogues' maintenance and extension process

Publish RWD & metadata catalogue good practice guide

Stakeholder engagement to populate catalogues, incl. with patients organisations

Consult on reflection paper on Patient Experience data

Engage with ERNs/registries

Review utility of mobile Health data

Review utility of social media

EU Network skills

Deliver Big Data curricula training through EUNTC (V)

Offer training modules to patients, healthcare professionals & academics (v-)

Adopt genomics data curriculum

Review training curricula (v-)

Dialogue with patients' organisations on training needs

Survey EMRN training needs (V)

| | Q1 - Q3 | Q4 | Q1 | Q2 | Q3 | Q4 | Q1 |

*Scope applicable to veterinary aspects ((v-) if tbc)

Timeline: 2023 (Q1–Q3, Q4) · 2024 (Q1, Q2, Q3, Q4) · 2025 (Q1)

EU Network processes
- Publish use cases
- Develop portfolio of use cases for use by EMRN
- Publish report on RWE framework to support regulatory decision-making
- Publish review of RWE pilots
- Publish EU network RWE processes overview (v-)
- Publish official glossary (v-)
- Develop use cases for Patient experience data
- Develop use case for genomics data
- Publish report of RWE in regulatory decision-making (v-)
- Share good practices on regulatory data science, management and software

Network capability to analyse
- CHMP clinical trial raw data pilot
- CT raw data interim report
- CT raw data final report
- POC of nonclinical raw data analysis
- Review on nonclinical raw data POC
- Scope knowledge sharing fora on advance analytics
- BDSG discussion on CMC data analysis (v-)
- Publish Artificial Intelligence workplan to 2028 (V)
- Deliver enhanced EV data analytics
- Launch new EV website
- Review EMRN computing capability to analyse Big Data (v-)
- Deliver recommendations on capability and capacity (v-)
- Experimentation of advanced analytics, including AI
- Release of AI knowledge management tool for core regulatory processes

Delivery of expert advice
- Methodology ESEC launch
- Publish data and methods guidance roadmap (v-)
- Update roadmap (v-)
- Consolidate Methodology WP
- Establish AI and RWE Specialised Interest Area of the ESEC (v-)
- Establish CT Raw data Cluster of Excellence
- Consult on AI reflection paper
- Consult on RWE reflection paper
- Publish final AI reflection paper (v-)

Governance framework
- Revised BDSG mandate adoption (V)
- Support EHDS and TEHDAS II
- Assess EHDS impact on medicines regulation
- BDSG recommendation on EHDS implementation
- Support revised Pharmaceutical Strategy for Europe
- Modular delivery of data protection training (V)
- Development of EMRN data strategy (V)
- Publish EMRN data strategy (V)

*Scope applicable to veterinary aspects ((v-) if tbc)

International initiatives

- International collaboration on framework for RWE
- ○ ICH new topic proposal on RWE
- ○ Consider new ICH topic, e.g. data quality, RWD data standards
- Consult on ICH reflection paper on RWE terminologies and studies
- ICH concept paper on RWE
- Consult on ICH M14 Use of RWD for safety assessment of medicines
- Final ICH M14 guideline
- Continue RWE global regulators cooperation under the umbrella of ICMRA

Stakeholder engagement

- Industry meetings (V)
- Industry meetings (V)
- Big data forum (V)
- Industry meetings (V)
- Industry meetings (V)
- Big data forum (V)
- Workshop on CT analytics
- Workshop on RWD quality & experience in use of RWE
- Workshop on AI
- Workshop on ADR data quality
- Workshop on RWE methodologies/research
- Workshop on AI
- Workshop on CT Raw data
- Workshop on registries
- Workshop on Pharmacogenomics
- Develop and implement Network change management (v-)

Veterinary recommendations (V)

- Big data forum (V)
- Big data forum (V)
- EU Veterinary Big Data Workplan to 2025 adoption (V)
- EU Veterinary data hub establishment (V)

Timeline: 2023 Q1 - Q3 | Q4 | 2024 Q1 | Q2 | Q3 | Q4 | 2025 Q1

Topic description

DARWIN EU®

The Data Analysis and Real World Interrogation Network (DARWIN EU®) is a federated network to enable access and analysis of real-world data (RWD).

Following the establishment of the DARWIN EU Coordination Centre in 2022, implementation activities continue with a focus on onboarding additional data partners and initiating studies every year, gradually anchoring the methodology and results of RWD analyses in regulatory decision-making at EMA and the European Medicines Regulatory Network (EMRN).

Pilot activities are expanding to support Member States national use cases, the European Health Data Space (EHDS) and to cooperate with the European Centre for Disease Prevention and Control (ECDC) and bodies responsible for Health Technology Assessments (HTA) as well as payers.

For more information please check out our website:

Key dates:

- Q4 2022 – 2024 Pilot RWE use with HTA & payers

- 2023 Phased routine access to RWE for EMRN: COMP & PDCO

- 2023 - 2024 Support RWD analysis pilot through EHDS

- Q4 2023 ≥ 16 studies per year and 10 additional data partners

- 2023 – 2024 Pilot RWE use with ECDC incl. in health crisis

- 2024 Phased routine access to RWE for EMRN: CHMP & CAT

- Q4 2024 EHDS2 pilot interim report

- 2024 – 2025 Pilot use cases with NCAs

Data quality & representativeness

To inform the selection of data for studies and judge the evidentiary value of studies, data quality should be understood and efforts should be made for quality to be increased.

Engagement with stakeholders and leveraging the work of external parties remain critical to delivering on data quality and representativeness.

Following the publication of the final version of the data quality framework for medicines regulation, specific quality chapters will be established for new data domains, starting with RWD and ADR.

The work will also continue in the coming years to strengthen the EMA data qualifications process and through further collaboration with the proposed EHDS.

Key dates:

- Q2 2023 - 2025 Support EHDS data quality & labelling work

- Q3 2023 Publish final data quality framework v.1. (V)

- Q1 2024 Prioritisation of next domains for data quality considerations (v-)

- Q1 2024 Advice on revised data qualification process

- Q3 2024 Consult on RWD quality chapter

- Q4 2024 Publish RWD quality chapter

- Q4 2024 Consult on ADR quality chapter

- Q1 2025 Publish ADR quality chapter

Data discoverability

Closely linked to the work on data quality, to be able to include novel data sources as the basis for studies that support regulatory decision-making, the improvement of their discoverability is critical.

After the publication of the [first metadata list for real-world data sources and studies](#), the EMA-HMA catalogues of data sources and non-interventional studies will be launched, supported by a good practice guide.

Engagement with stakeholders and, in particular with European Reference Networks (ERNs), and work on registries will be strengthened, also helping to populate the catalogues.

Looking forward to the use of new data type, the utility of mobile Health and social media as data sources will be explored and a public consultation on a reflection paper on patient experience data will be launched.

Key dates:

- Q4 2023 Publish metadata & catalogue good practice guide

- Q1 2024 Launch EMA-HMA catalogues of data sources and non-interventional studies

- 2024 Stakeholder engagement to populate catalogues, incl. with patients organisations

- Q1 2024 Engage with ERNs/registries

- Q2 2024 Review utility of mobile Health data

- Q3 2024 Review utility of social media

- Q2/Q3 2024 Review RWD and metadata catalogues' maintenance and extension process

- Q3 2024 Consult on reflection paper on patient experience data

- Q1 2025 Extend RWD and metadata catalogues to veterinary data sources (V)

EU Network skills

To support the development of an expert workforce in the Network able to advise on and interpret big data, the agreed training curricula on biostatistics, pharmacoepidemiology/real world evidence and data science will be further developed with content rolled out to the EU Network via the EU Network Training Centre, with targeted access provided to patients, healthcare professionals and academics.

A curriculum on genomics data will be developed.

Skills will be further enhanced on the basis of a review of training delivery, adequacy as well as identification of further needs, including the needs of patients' organisations.

Key dates:

- 2023 – 2025 Deliver Big Data curricula training through EUNTC (V)

- Q4 2023 Review training curricula (v-)

- 2024 Survey EMRN training needs (V)

- Q1 2024 Offer training modules to patients, healthcare professionals & academics (v-)

- Q1 2024 Dialogue with patients' organisations on training needs

- Q4 2024 Adopt genomics data curriculum

EU Network processes

Sharing use cases and learning from pilots, good practices on regulatory data science, management and software will contribute to process improvements, guidance development and uptake of real-world evidence (RWE) by the EMRN for better regulatory decision-making including greater preparedness for health crisis response.

Reports on pilot studies on the use of RWE by the EU Network will be published incrementally and concluded in 2025 with an overview report on RWE in regulatory decision-making.

A portfolio of RWE use cases will be published and new uses cases for genomics and patient experience data will be explored.

Key dates:

- 2023 – 2025 Develop portfolio of use cases for use by EMRN

- 2023 – 2025 Share good practices on regulatory data science, management and software

- Q2 2023 Publish report on RWE framework to support regulatory decision-making

- Q4 2023 Publish use cases

- Q1 2024 Publish EU network RWE processes overview (v-)

- Q2 2024 Publish review of RWE pilots

- Q2 2024 Publish official glossary (v-)

- Q4 2024 Develop use case for genomics data

- Q4 2024 Develop Use case for Patient experience data

- Q4 2024 Revise use cases

- Q1 2025 Publish report of RWE in regulatory decision-making (v-)

Network capability to analyse

The EMRN computing capacity will be reviewed and options for knowledge sharing within EMRN will be discussed to support the network capabilities.

Analytics initiatives will include:

- The CHMP clinical trial raw data pilot will continue to clarify the benefits and practicalities of access to individual patient data from clinical trials and will help the EU medicines regulatory network to make an informed decision in 2025 on the place of raw data.

- A proof on concept on the analysis of nonclinical raw data will explore how to improve the quality and efficiency of assessment procedures.

Key dates:

- 2022-2024 CHMP clinical trial raw data pilot

- Q3 2023 Scope advanced analytics knowledge sharing fora

- Q2/Q3 2023 Review EMRN computing capability

- Q4 2023 Publish Artificial Intelligence workplan to 2028 (V)

- 2024 – 2025 Experimentation of advanced analytics, incl. AI

- Q1 2024 Release of AI knowledge management tool for core regulatory processes

- 2024 Deliver recommendations on capability and capacity (v-)

- Q2 2024 - 2025 Deliver enhanced EV data analysis

- Q2 2024 CT raw data interim report

- Q4 2024 CT raw data pilot final report

- The data analysis for assessment and inspection of manufacturing data will be discussed, while safety monitoring will benefit from enhanced EudraVigilance data analytics.

- Experimentation of advanced analytics, including AI and Large Language Models will intensify in collaboration with the Network. AI tools to mine knowledge from regulatory documents will be rolled out starting with scientific advice opinions and MAAs.

The Artificial Intelligence workplan to 2028 (V) will set out a collaborative and coordinated strategy to maximise the benefits of AI to stakeholders and ensure EMRN remains at the forefront in benefiting from AI in medicines regulation.

- Q4 2024 BDSG discussion on CMC data analysis (v-)
- 2024 – 2025 POC of nonclinical raw data analysis
- Q1 2025 Launch new EudraVigilance website
- Q1 2025 Review on nonclinical raw data POC

Delivery of expert advice

Expert advice is needed to deliver robust assessment and decision-making by EMA regulatory committees.

Activities under the 2023-2025 work plan will follow the roadmap for the development of guidance across data and methods and will leverage the newly established EMA Methodology Working Party.

In particular, the work will include the publication of reflection papers on artificial intelligence (AI) for use in regulatory processes and on RWE and will lead to further development of guidance in the future.

Expert advice will also be strengthened through the Methodology European Specialised Expert Community (ESEC) with two new specialised interest areas in AI & RWE. Collaboration in analysis of raw data from clinical trials will be further fostered through the establishment of a cluster of excellence.

Key dates:

- Q4 2022 - 2023 Consolidate Methodology WP

- Q2 2023 Methodology ESEC launch

- Q3 – Q4 2023 Consult on reflection paper on AI (V)

- Q4 2023 Publish data and methods guidance roadmap (v-)

- 2023 - 2025 Establish AI and RWE Specialised Interest Area of the ESEC (v-)

- 2023 - 2025 Establish CT Raw data Cluster of Excellence

- Q2 2024 Consult on RWE reflection paper

- Q3 2024 Publish final AI reflection paper (V)

- Q4 2024 Update roadmap (v-)

Governance framework

The [BDSG mandate](#) and Network Data Board mandates have been revised in 2023 to improve data governance in the EMRN.

To ensure alignment, collaboration and preparedness for a changing policy environment, the BDSG will continue to support the Commission and its joint action TEHDAS in the preparations for the future EHDS and provide recommendation on the impact assessment of EHDS on medicines regulation. BDSG will also support discussion on the revised Pharmaceutical Strategy for Europe.

Introduction training on data protection principles followed by specialised topics will be explored and gradually delivered.

Key dates:

- Q4 2022 – 2025 Modular delivery of data protection training (V)

- Q4 2022 Revised BDSG mandate adoption (V)

- 2023 - 2024 Assess EHDS impact on medicines regulation

- 2023 – 2024 Support EHDS and TEHDAS II

- 2023 – 2024 Support revised Pharmaceutical Strategy for Europe

- 2024 Development of EMRN data strategy (V)

- Q4 2024 BDSG recommendation on EHDS implementation

- Q4 2024 Publish final EMRN data strategy (V)

International initiatives

Convergence with international partners on Big Data including guidelines will enable efficient and effective evidence generation by stakeholders.

Following the [ICMRA summit with international regulators in 2022](), the international collaboration on RWE will continue throughout the workplan.

Collaboration at ICH will also intensify with the launch of two public consultations, on the ICH reflection paper on proposed international harmonisation of RWE terminology and convergence of general principles regarding planning and reporting of studies using RWD, and on the draft ICH M14 guideline on the use of RWD for safety assessment of medicines. This will be followed by further guideline development and consideration on new ICH topics, e.g. data quality, RWD data standards.

Key dates:

- Q4 2022 – 2025: International collaboration on framework for RWE

- 2023-2025: Continue RWE global regulators cooperation under the umbrella of ICMRA

- Q3 2023: Consult on ICH reflection paper on RWE terminologies and studies

- Q3 2024: Consult on ICH M14 Use of RWD for safety assessment of medicines

- Q4 2024: Final ICH M14 guideline

- Q4 2024: ICH new topic proposal on RWE

- Q1 2025: Consider new ICH topic, e.g., data quality, RWD data standards

- 2025: ICH concept paper on RWE

EU BD stakeholder engagement

Listening to stakeholders, increased transparency and adequate communication will be critical to build trust in what is delivered and maximize the network transformation to data-driven regulation.

The Big Data stakeholder forum will continue to be organised annually and will be complemented by topic specific meetings and workshops held throughout the period of this work plan.

Biannual industry meetings will be held, and a network change management strategy on Big Data will be developed and implemented.

Key dates:

- 2023 – 2025 Biannual industry meeting (V)

- 2023 – 2025 Yearly multistakeholder forum (V)

- 2024 - 2025 Develop and implement Network change management
  strategy (v-)

- Q2 2023 Workshop on RWD quality and experience in use of RWE

- Q4 2023 AI workshop

- Q1 2024 Workshop on CT analytics

- Q1 2024 Workshop on registries

- Q1 2024 Workshop on ADR data quality

- Q2 2024 Workshop on RWE methodologies/research

- Q3 2024 Workshop on Pharmacogenomics

- Q4 2024 AI workshop

- Q1 2025 Workshop on CT Raw data

Veterinary recommendations

Synergies in the use of data between the human and veterinary domains can catalyse the network transformation to a data-driven regulation.

Following the adoption of the [EU Veterinary Big Data strategy](#) in 2022, the [EU Veterinary Big Data Workplan to 2025](#) will deliver the implementation of the strategy during the work period.

The Veterinary Big Data stakeholder forum will be organised annually and be complemented by ad-hoc workshops on key business areas to exchange views and experience with stakeholders.

Key dates:

- Q4 2022 2nd Veterinary Big Data Stakeholder Forum
- Q2 2023 EU Veterinary Big Data Workplan to 2025 adoption
- Q2 2023 EU Veterinary data hub establishment
- Q4 2023 3rd Veterinary Big Data Stakeholder Forum
- Q4 2024 4th Veterinary Big Data Stakeholder Forum

Annexes

Priority recommendations of the HMA-EMA joint big data task force*

DARWIN EU

Data Analysis and Real World Interrogation Network - DARWIN. Build the business case with stakeholders and secure funding to establish and maintain a secure EU data platform that supports better decision-making on medicines by informing those decisions with robust evidence from healthcare.

Data quality & representativeness

Establish an EU framework for data quality and representativeness. Develop guidelines, a strengthened process for data qualification through scientific advice, and promote across Member States the uptake of electronic health records, registries, genomics data, and secure data availability.

Data discoverability

Identify key metadata for regulatory decision-making on the choice of data source, strengthen the current ENCePP resources database to signpost to the most appropriate data, and promote the use of the FAIR principles (Findable, Accessible, Interoperable and Reusable).

EU Network skills

Develop a big data training curriculum and strategy based on a skills analysis across the network, collaborate with external experts including academia, and target recruitment of data scientists, omics specialists, biostatisticians, epidemiologists, and experts in advanced analytics and AI.

EU Network processes

Launch a 'big data learnings initiative' where submissions that include big data are tracked and outcomes reviewed, with learnings fed into reflection papers and guidelines. Enhance the existing EU PAS register to increase transparency on study methods.

Network capability to analyse

Build computing capacity to receive, store, manage and analyse large data sets including patient level data (PLD), establish a network of analytics centres linked to regulatory agencies, and strengthen the network's ability to validate AI algorithms.

Delivery of expert advice

Build on the existing working party structure to establish a Methodologies Working Party that encompasses biostatistics, modelling and simulation, extrapolation, pharmacokinetics, real world data, epidemiology and advanced analytics, and establish an Omics Working Party that builds on and reinforces the existing pharmacogenomics group.

Governance framework

Engage with initiatives on the implementation of EU data protection regulations to deliver data protection by design, engage with patients and healthcare professionals on data governance, and establish an Ethics Advisory Committee.

International initiatives

Support the development of guidelines at international multilateral fora, a data standardisation strategy delivered through standards bodies, and bilateral collaboration and sharing of best practice with international partners.

Stakeholder engagement

Dialogue actively with key EU stakeholders, including patients, healthcare professionals, industry, HTA bodies, payers, device regulators and technology companies. Establish key communication points in each agency and build a resource of key messages and communication materials on regulation and big data.

* After these recommendations were made by the HMA-EMA joint Big Data Task Force, veterinary recommendations have been included in the scope of the Big Data Steering Group

European Medicines Agency