HMA
Heads of Medicines Agencies

EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

Big Data Steering Group

Big Data Workplan 2022-2025

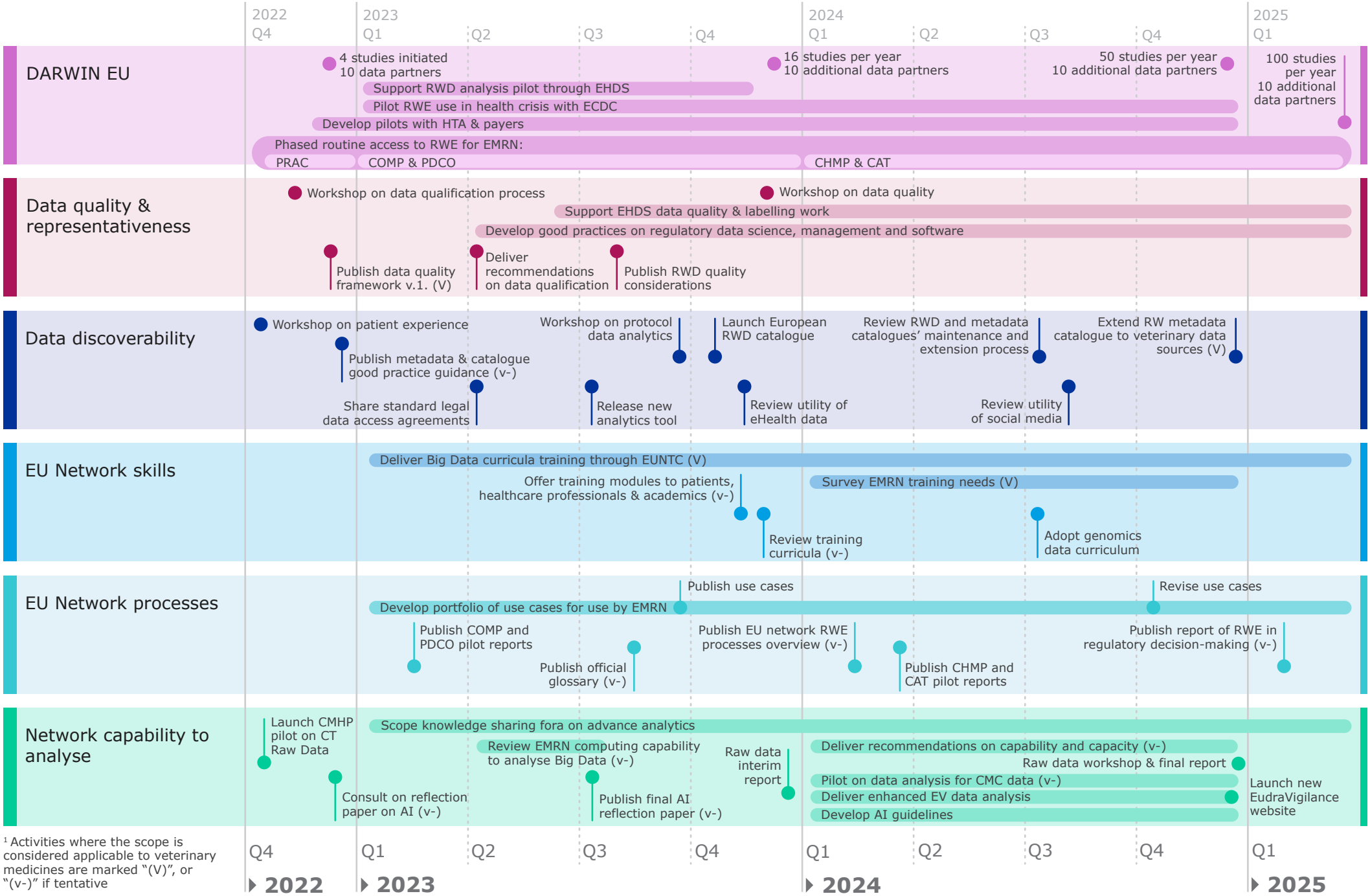
The vision on Big Data is a **strengthened regulatory system that can efficiently integrate data analysis into its assessment processes** to improve decision making.

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 3rd joint HMA-EMA Big Data Steering Group (BDSG) workplan was adopted in July 2022 and covers 2022 to 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 appropriate. Activities relevant only to veterinary medicines are included in a dedicated section.

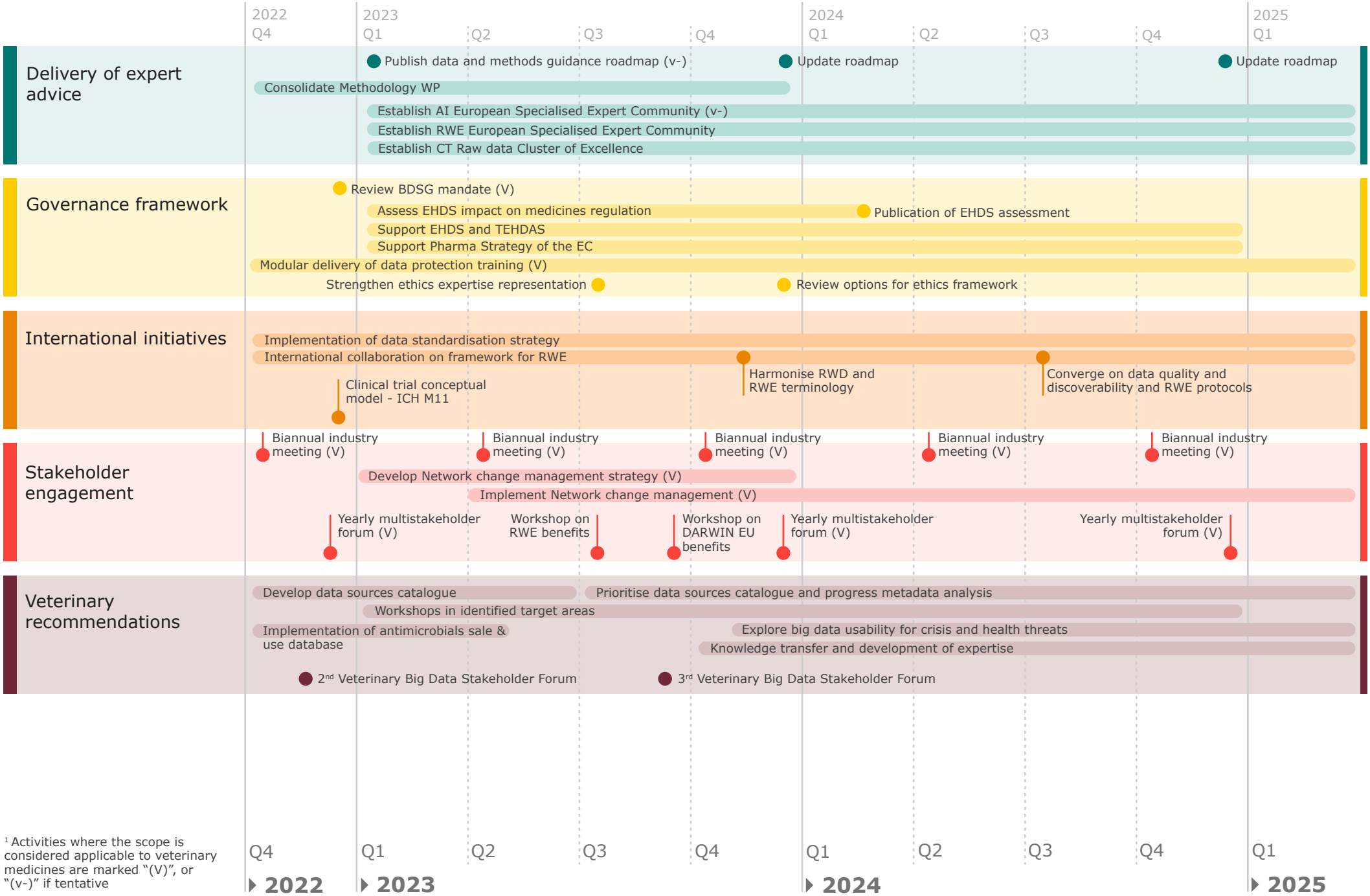
Content:

THE HMA-EMA JOINT BIG DATA STEERING GROUP WORKPLAN [1]



[1] Activities where the scope is considered applicable to veterinary medicines are marked "(V)", or "(v-)" if tentative

THE HMA-EMA JOINT BIG DATA STEERING GROUP WORKPLAN [1]



[1] Activities where the scope is considered applicable to veterinary medicines are marked "(V)", or "(v-)" if tentative

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 start with a focus on onboarding data partners and initiating studies, gradually anchoring the methodology and results of RWD analyses in regulatory decision-making at EMA and the European Medicines Regulatory Network (EMRN). Pilot activities will expand to support 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.

Q4 2022	4 studies initiated and 10 data partners
Q4 2022	Phased routine access to RWE for EMRN: PRAC
Q4 2022 - 2024	Develop pilots with HTA & payers
2023	Phased routine access to RWE for EMRN: COMP & PDCO
2023	Support RWD analysis pilot through EHDS
Q4 2023	16 studies per year and 10 additional data partners
2023 - 2024	Pilot RWE use in health crisis with ECDC
2024	Phased routine access to RWE for EMRN: CHMP & CAT
Q4 2024	50 studies per year and 10 additional data partners
Q1 2025	100 studies per year and 10 additional data partners

Data quality & representativeness

Engagement with stakeholders and leveraging the work of external parties remain critical to delivering on data quality and representativeness. Therefore, collaboration will continue with the joint action 'Towards A European Health Data Space – TEHDAS' dedicated to the technical and scientific aspects of data quality. Following the analysis and exchanges on data quality with a wide range of stakeholders the first version of a data quality framework for the EU Regulatory Network will be delivered in late 2022, followed by work in the coming years to strengthen the EMA data qualifications process and further collaboration with the EHDS.

Q4 2022	Workshop on data qualification process
Q4 2022	Publish data quality framework v.1. (V)
Q2 2023	Deliver recommendations on data qualification
Q2 2023 - 2025	Develop good practices on regulatory data science, management and software
Q2 2023 - 2025	Support EHDS data quality & labelling work
Q3 2023	Publish RWD quality considerations
Q4 2023	Workshop on data quality

Data discoverability

Closely linked to the work on data quality is the agreement on metadata to describe and identify RWD sets. The publication of a good practice guide on metadata will be followed by a public catalogue of European RWD. Establishing the capability to find information from structured and unstructured regulatory documents will be enhanced through the development of analytics tools and the introduction of standardised study protocols. Furthermore, the utility of eHealth and social media as data sources will be considered.

Q4 2022	Workshop on patient experience
Q4 2022	Publish metadata & catalogue good practice guidance (v-)
Q2 2023	Share standard legal data access agreements
Q3 2023	Release new analytics tool
Q3 2023	Workshop on protocol data analytics
Q4 2023	Launch European RWD catalogue
Q4 2023	Review utility of eHealth data
Q3 2024	Review utility of social media

Q3 2024	Review RWD and metadata catalogues' maintenance and extension process
Q4 2024	Extend RW metadata catalogue to veterinary data sources (V)

EU Network skills

Agreed training curricula on biostatistics, pharmacoepidemiology and data science will be further developed and delivered with selected modules opened to patients, healthcare professionals and academics. A curriculum on genomics data will be adopted. Skills will be further enhanced on the basis of a review of training delivery, adequacy and further needs.

2023 - 2025	Deliver Big Data curricula training through EUNTC (V)
Q4 2023	Offer training modules to patients, healthcare professionals & academics (v-)
Q4 2023	Review training curricula (v-)
2024	Survey EMRN training needs (V)
Q3 2024	Adopt genomics data curriculum

EU Network processes

Reports on pilot studies on the use of real world evidence (RWE) by EMA scientific committees will be published incrementally and concluded in 2025 with a report on RWE in regulatory decision-making. A portfolio of RWE use cases will be published to support uptake of RWE by the ERMN.

2023 - 2025	Develop portfolio of use cases for use by EMRN
Q1 2023	Publish COMP and PDCO pilot reports
Q3 2023	Publish official glossary (v-)
Q3 2023	Publish use cases
Q1 2024	Publish EU network RWE processes overview (v-)
Q1 2024	Publish CHMP and CAT pilot reports
Q4 2024	Revise use cases
Q1 2025	Publish report of RWE in regulatory decision-making (v-)

Network capability to analyse

A proof-of-concept pilot will be performed to clarify the benefits and practicalities of access to individual patient data from clinical trials (raw data) in the scientific assessment of medicines. Learnings from the pilot will help the EU medicines regulatory network to make an informed decision on the place of raw data in regulatory decision-making. Data analysis for assessment and inspection of manufacturing data will be explored, while safety monitoring will benefit from enhanced EudraVigilance data analytics. The EMRN capacity and capability for computing will be reviewed and actions taken based on recommendations. A reflection paper on artificial intelligence (AI) for use in regulatory processes may lead to development of guidelines. Options for knowledge sharing within EMRN will also be explored to support network capabilities.

Q4 2022	Launch CHMP pilot on CT Raw Data
Q4 2022	Consult on reflection paper on AI (v-)
2023 - 2025	Scope knowledge sharing fora on advance analytics
Q2 - Q3 2023	Review EMRN computing capability to analyse Big Data (v-)
Q3 2023	Publish final AI reflection paper (v-)
Q4 2023	Raw data interim report
2024	Deliver recommendations on capability and capacity (v-)
2024	Pilot on data analysis for CMC data (v-)
2024	Deliver enhanced EV data analysis
2024	Develop AI guidelines
Q4 2024	Raw data workshop & final report
Q4 2024	Launch new EudraVigilance website

Delivery of expert advice

Activities under the 2022-2025 work plan will follow the roadmap for the development of guidance across data and methods and will leverage the new EMA Methodologies Working Party. The work includes development of guidance in particular concerning RWE. Expert advice will be strengthened through European Specialised Expert Communities (ESEC) initially in AI & RWE. Collaboration in analysis of raw data from clinical trials will be further fostered through the establishment of a cluster of excellence.

Q4 2022 - 2023	Consolidate Methodology WP
Q1 2023	Publish data and methods guidance roadmap (v-)
2023 - 2025	Establish AI European Specialised Expert Community (v-)
2023 - 2025	Establish RWE European Specialised Expert Community
2023 - 2025	Establish CT Raw data Cluster of Excellence
Q4 2023	Update roadmap
Q4 2024	Update roadmap

Governance framework

The BDSG and EU Network Data Board mandates will be reviewed in late 2022 to improve data governance in the EMRN. To ensure alignment, collaboration and preparedness the BDSG will continue to support TEHDAS in the preparations for the future EHDS and a provisional assessment of the impact of EHDS on medicines regulation will be performed. Introduction training on data protection principles followed by specialised topics will be delivered.
The representation of expertise in ethics will be strengthened and strengthening the framework for ethics in data analysis will be explored.

Q4 2022 - 2025	Modular delivery of data protection training (V)
Q4 2022	Review BDSG mandate (V)
2023 - Q1 2024	Assess EHDS impact on medicines regulation
2023 - 2024	Support EHDS and TEHDAS
2023 - 2024	Support Pharma Strategy of the EC
Q3 2023	Strengthen ethics expertise representation
Q4 2023	Review options for ethics framework
Q1 2024	Publication of EHDS assessment

International initiatives

The implementation of the EMRN data standardisation strategy for medicines regulation will advance during the work plan period. The international collaboration on RWE will be intensified, catalysed by the summit with international regulators in 2022.

Q4 2022	Clinical trial conceptual model - ICH M11
Q4 2022 - 2025	Implementation of data standardisation strategy
Q4 2022 - 2025	International collaboration on framework for RWE
Q4 2023	Harmonise RWD and RWE terminology
Q3 2024	Converge on data quality and discoverability and RWE protocols

Stakeholder engagement

The Big Data stakeholder forum will be organised annually and be complemented by topic specific meetings and workshops held throughout the period of this work plan. Biannual industry meetings will be held. A network change management strategy on Big Data will be developed and implemented.

Q4 2022	Biannual industry meeting (V)
Q4 2022	Yearly multistakeholder forum (V)
2023	Develop Network change management strategy (V)
Q2 2023 - 2025	Implement Network change management (V)
Q2 2023	Biannual industry meeting (V)
Q3 2023	Workshop on RWE benefits
Q3 2023	Workshop on DARWIN EU benefits
Q4 2023	Biannual industry meeting (V)
Q4 2023	Yearly multistakeholder forum (V)

Q2 2024	Biannual industry meeting (V)
Q4 2024	Biannual industry meeting (V)
Q4 2024	Yearly multistakeholder forum (V)

Veterinary recommendations

The EU Veterinary Big Data strategy was adopted in 2022 and will be implemented during the work period. This includes delivery of a catalogue of veterinary data sources, finalising the system envisaged by the VMP Regulation on antimicrobials sale and use data and evolution of training curricula. Ad-hoc workshops on key business areas will be organised to exchange views and experience with stakeholders.

Q4 2022	2nd Veterinary Big Data Stakeholder Forum
2022 - Q2 2023	Develop data sources catalogue
2022 - Q2 2023	Implementation of antimicrobials sale & use database
2023 - 2024	Workshops in identified areas
Q3 2023 - 2025	Prioritise data sources catalogue and progress metadata analysis
Q3 2023	3rd Veterinary Big Data Stakeholder Forum
Q4 2023 - 2025	Explore big data usability for crisis and health threats
Q4 2023 - 2025	Knowledge transfer and development of expertise

ANNEX I : PRIORITY RECOMMENDATIONS OF THE HMA-EMA JOINT BIG DATA TASK FORCE [2]

I	Deliver a sustainable platform to access and analyse healthcare data from across the 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.
II	Establish an EU framework for data quality and 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.
III	Enable 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).
IV	Develop EU network skills in big data	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.
V	Strengthen EU network processes for big data submissions	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.
VI	Build EU Network capability to analyse big data	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.
VII	Modernise the 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.
VIII	Ensure data are managed and analysed within a secure and ethical 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.
IX	Collaborate with international initiatives on big data.	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.
X	Create an EU big data 'stakeholder implementation forum'	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.

[2] 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.