HMA
Heads of Medicines Agencies

EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

Big Data Steering Group

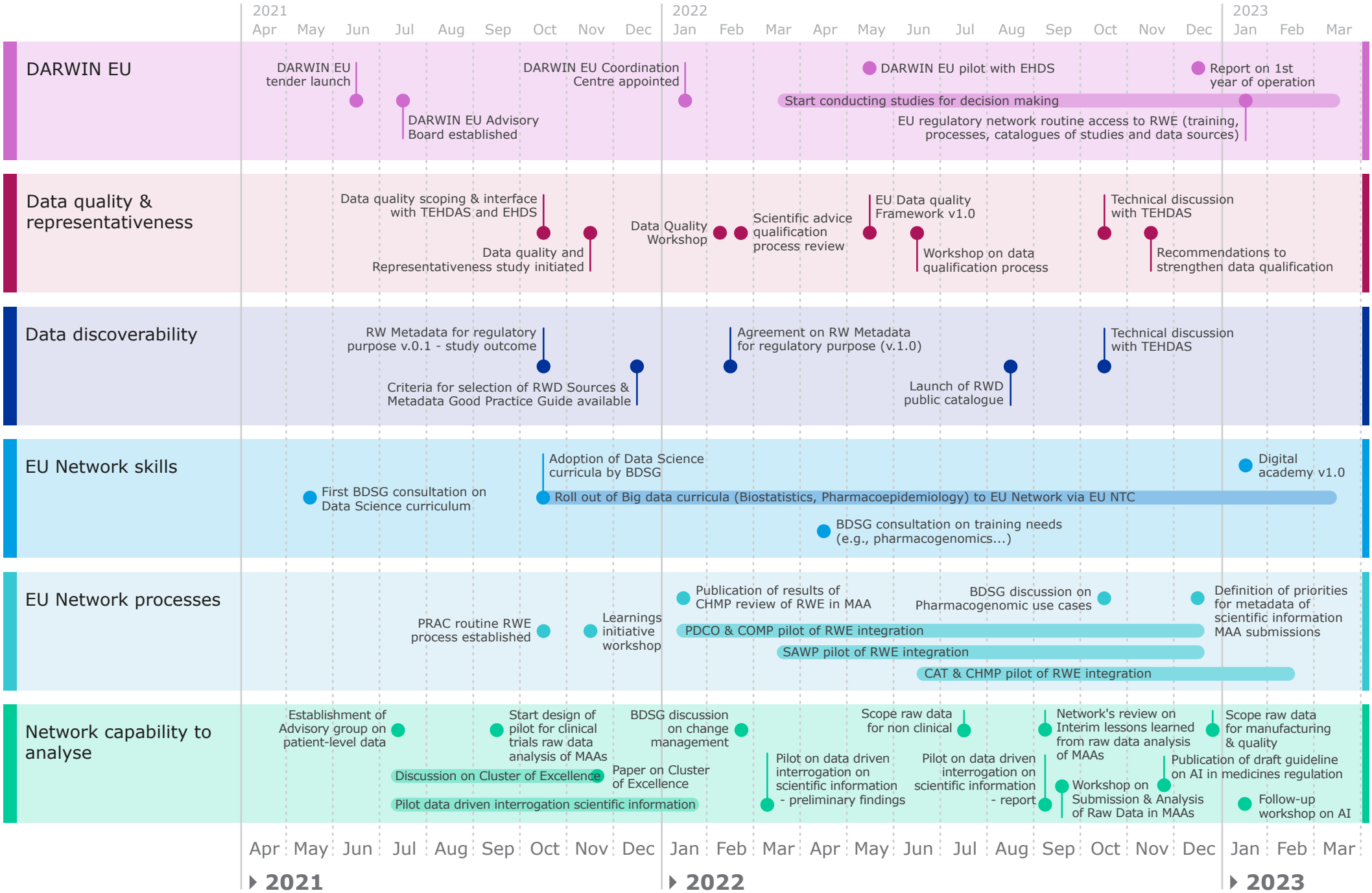
Workplan 2021-2023

The 2021-2023 HMA-EMA joint Big Data Steering Group (BDSG) workplan was adopted on **18 June 2021.** This document introduces each topic and outlines key deliverables. The plan was prepared based on the BDSG mandate, the continuation of the activities launched in 2020-21 and the need to address new topics. The document is structured in line with the key recommendations of the Big Data Task Force (see Annex I). Information security and ethical data governance are at the core of the work of the BDSG.

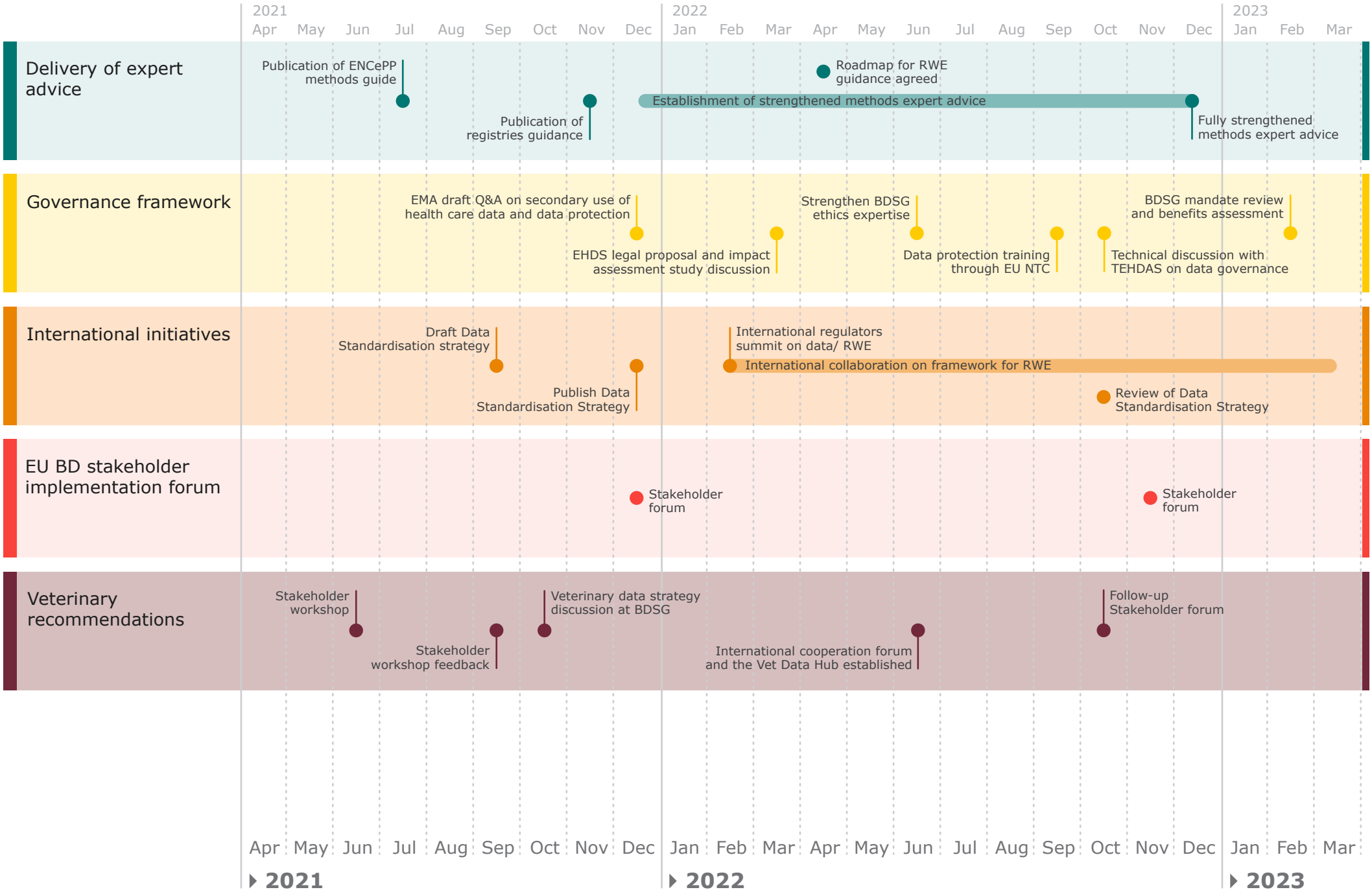
Implementation of the actions in the Big Data Steering Group workplan 2021-23 will need to be flexible given uncertainties on the resources that will need to be prioritised towards the regulatory response to the COVID-19 pandemic.

Content:

THE HMA-EMA JOINT BIG DATA STEERING GROUP WORKPLAN



THE HMA-EMA JOINT BIG DATA STEERING GROUP WORKPLAN



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 2021 establishment of the DARWIN EU Advisory Board and the launch of the tender to select the DARWIN EU Coordination Centre, implementation work will start in early 2022. The Coordination Centre will be appointed and will start to assemble the DARWIN EU network on behalf of the EU Regulatory network. The first studies will be delivered to test tools and processes and to deliver early benefits to EMA committees and EU patients. Through pilots, the DARWIN EU network will test connection to the planned European Health Data Space (EHDS).

Jun 21	DARWIN EU tender launch
Jul 21	DARWIN EU Advisory Board established
Jan 22	DARWIN EU Coordination Centre appointed
Mar 22	Start conducting studies for decision making
May 22	DARWIN EU pilot with EHDS
Dec 22	Report on first year of operation
Jan 23	EU regulatory network routine access to RWE (training, processes, catalogues of studies and data sources)

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' focussed on the technical and scientific aspects of data quality. In 2021 an external study was launched that will analyse existing data quality initiatives and discuss data quality with a wide range of stakeholders.

In 2022, based on this work, the first version of the data quality framework for the EU Regulatory Network will be delivered. The Scientific Advice data qualification process will be reviewed starting with a stakeholder workshop in 2022.

Oct 21	Data quality scoping & interface with TEHDAS and EHDS
Nov 21	Data quality & Representativeness study initiated
Feb 22	Data Quality Workshop
Feb 22	Scientific advice qualification process review
May 22	EU Data Quality Framework v1.0
Jun 22	Workshop on data qualification process
Oct 22	Technical discussion with TEHDAS
Nov 22	Recommendations to strengthen data qualification

Data discoverability

Closely linked to the work on data quality is the agreement on metadata to be used to describe and identify RWD sets. Following the 2021 preparatory work including an external study, in 2022 criteria for the selection of RWD sources, a metadata good practice guide and a public catalogue of European RWD will be launched. This work will link to the TEHDAS work programme and will support the identification of relevant data for studies and the interpretation of results submitted for regulatory decision-making, including studies through the DARWIN EU network.

Oct 21	RW Metadata for regulatory purpose v.0.1 - study outcome
Dec 21	Criteria for selection of RWD Sources & Metadata Good Practice Guide available
Feb 22	Agreement on RW Metadata for regulatory purpose (v.1.0)
Aug 22	Launch of the RWD public catalogue
Oct 22	Technical discussion with TEHDAS

EU Network skills

Following the adoption of training curricula on Biostatistics and Pharmacoepidemiology, a Data Science curriculum will be adopted in 2021. In 2022, the delivery of training will be rolled-out to the EU Regulatory Network including through the engagement of external training partners.
Work will continue to improve the existing curricula and to identify new training needs guided by the results of the 2021 survey of EU Network skills.

May 21	First BDSG consultation on Data Science curriculum
Oct 21	Roll out of Big Data curricula (Biostatistics, Pharmacoepidemiology) to EU Network via EU NTC
Oct 21	Adoption of the Data Science curricula by the BDSG
Apr 22	BDSG consultation on training needs (e.g., pharmacogenomics...)
Jan 23	Digital academy v1.0

EU Network processes

Critical to the development of processes, to guidance for industry, and to delivery of data-driven decisions will be to systematically learn from applications to the Network that include Big Data. In late 2021 a 'learnings initiative' workshop will be held. This workshop will include the results of the CHMP review of Real World Evidence (RWE) in Marketing Authorisation Applications (MAA) and extensions of indications and past piloting of RWD analysis in committee decision-making (notably with the PRAC). The workshop will inform process and guidance improvement in 2022-2023.
Starting from late 2021, PRAC will have established processes for RWD access to support its decision-making. In addition, RWD use cases will be confirmed and pilots will be run with the Scientific Advice Working Party (SAWP), the Paediatric Committee (PDCO), the Orphan Committee (COMP) and, later, the Committee on Advanced Therapies (CAT) and the Committee for Medicinal Products for Human use (CHMP).

By the end of this workplan, the processes for the delivery of real-world evidence to EMA committees will be established and its place in regulatory decision-making should, at least provisionally, be established.

Oct 21	PRAC routine RWE process established
Nov 21	Learnings initiative workshop
Jan 22	Publication of results of CHMP review of RWE in MAA
Jan 22	PDCO & COMP pilot of RWE integration
Mar 22	SAWP pilot of RWE integration
Jun 22	CAT & CHMP pilot of RWE integration
Oct 22	BDSG discussion on Pharmacogenomic use cases
Dec 22	Definition of priorities for metadata of scientific information MAA submissions

Network capability to analyse

Leveraging learnings from the 2021 pre-pilot and guided by a newly established advisory group on patient-level data, a full pilot of analysis of clinical trials raw data in MAAs will be conducted.
In 2022 stakeholders will be engaged in a workshop on clinical trials raw data. The business case and practicalities for analysis of raw data from non-clinical and from manufacturing and quality will be explored in the second half of 2022.
To complement the work on raw data, a pilot will be run on data driven interrogation of scientific information submitted as part of scientific advice requests and MAAs.
The BDSG will explore how data analysis clusters of excellence at national level can be fostered, including through mutual support and sharing of good practice.
Following the Artificial Intelligence (AI) workshop in 2021, work to draft a guideline on AI in medicines regulation will be developed through 2022 and a second workshop on AI will be held in early 2023, particularly to focus on building collaborations.

Jul 21	Discussion on Cluster of Excellence

Jul 21	Establishment of Advisory group on patient-level data
Jul 21	Pilot on data driven interrogation on scientific information
Sep 21	Start design of pilot for clinical trials raw data analysis of MAAs
Nov 21	Paper on Cluster of Excellence
Feb 22	BDSG discussion on change management
Mar 22	Pilot on data driven interrogation on scientific information - preliminary findings
Jul 22	Scope raw data for non clinical
Sep 22	Pilot on data driven interrogation on scientific information - report
Sep 22	Network's review on Interim lessons learned from raw data analysis of MAAs
Sep 22	Workshop on the Submission & Analysis of Raw Data in MAAs
Nov 22	Publication of draft guideline on AI in medicines regulation
Dec 22	Scope raw data for manufacturing and quality
Jan 23	Follow-up workshop on AI

Delivery of expert advice

Expert advice on Big Data and methodologies will be strengthened. Based on the EMA Management Board mandated model for the Agency's expert working parties, the gaps in Big Data expert advice will be addressed with modernised delivery of advice in 2022. Starting in 2021, methodology guidelines will be rolled-out starting with a substantial revision of the ENCePP guide to pharmacoepidemiological methods, new guidance on studies based on patient registries, and in 2022 a roadmap will be developed for the development of comprehensive guidance across data and methods (leveraging the new Methodologies Working Party).

Jul 21	Publication of ENCePP methods guide
Nov 21	Publication of registries guidance
Dec 21	Establishment of strengthened methods expert advice
Apr 22	Roadmap for RWE guidance agreed
Dec 22	Fully strengthened methods expert advice

Governance framework

To guide stakeholders, to support compliance and to enable public health research, a question and answer document on data protection in the context of secondary use of healthcare data will be finalised in late 2021 (subject to the publication of anticipated guidance from the European Data Protection Board).
Data protection training via the EU Network Training Centre will support the EU Regulatory Network from late 2022. Information security will continue to be a priority.
In 2022, the BDSG will strengthen its expertise on ethics ahead of a review of its mandate and an assessment of benefits that will be delivered in early 2023.
The BDSG will continue to prepare for the future EHDS and a technical workshop with TEHDAS on data governance is envisaged to ensure alignment, collaboration and preparedness.

Dec 21	EMA draft Q&A on secondary use of health care data and data protection
Mar 22	EHDS legal proposal and impact assessment study discussion
Jun 22	Strengthen BDSG ethics expertise
Sep 22	Data protection training through EU NTC
Oct 22	Technical discussion with TEHDAS on data governance
Feb 23	BDSG mandate review and benefits assessment

International initiatives

A data standardisation strategy for medicines regulation (including Big Data) will be agreed by the EU Regulatory Network and published in late 2021 with roll-out thereafter.
International collaboration on RWE will be intensified, catalysed by a summit with international regulators in 2022.

Sep 21	Draft Data Standardisation Strategy
Dec 21	Publish Data Standardisation Strategy
Feb 22	International regulators summit on data / RWE

Feb 22	International collaboration on framework for RWE
Oct 22	Review of Data Standardisation Strategy

EU BD stakeholder implementation forum

The big data stakeholder forum will continue with a plenary in late 2022 and again in 2023. These plenaries will be supplemented by topic specific meetings and workshops held throughout the period of this workplan.

Dec 21	Stakeholder forum
Nov 22	Stakeholder forum

Veterinary recommendations

The EMA Veterinary data strategy has been established in 2021. An international cooperation forum will be organised in 2022 and the Vet Data Hub will be launched, with a stakeholder workshop on veterinary data foreseen for late 2022.

Jun 21	Stakeholder workshop
Sep 21	Stakeholder workshop feedback
Oct 21	Veterinary data strategy discussion at BDSG
Jun 22	International cooperation forum and Vet Data Hub established
Oct 22	Follow-up Stakeholder forum

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

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.