18 January 2024
EMA/569310/2023

# Big Data Steering Group (BDSG): 2023 report

2023 marked significant progress in the transformation to data-driven regulation continued, in line with the Network Strategy to 2025 and BDSG workplan. The fourth BDSG workplan was published in July 2023 to progress the activities launched in 2020 and to address new topics.

This report is based on the third BDSG workplan and provides a summary of the key activities and achievements of the BDSG in 2023.

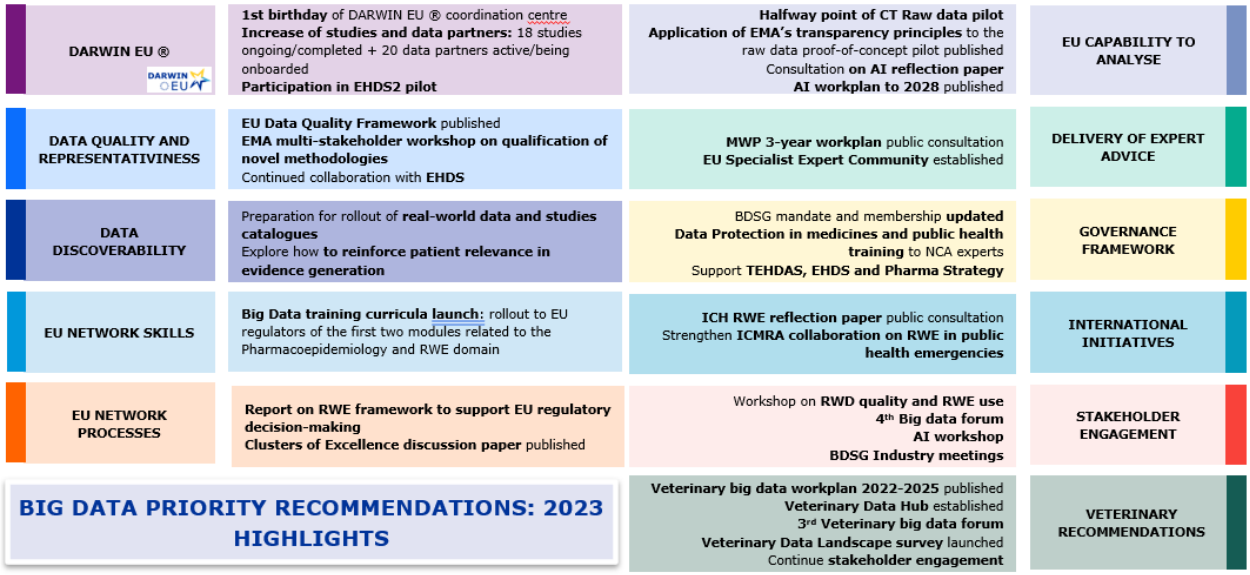
The Phase II report of the HMA-EMA joint Big Data Task Force (BDTF) and the proposal to establish a joint BDSG (superseding the Big Data Task Force) were endorsed by the Heads of Medicines Agencies (HMA) in November 2019 and EMA Management Board (EMA MB) in December 2019. The BDSG was established to advise HMA and EMA MB on the recommendations of the Big Data Task Force, covering human and veterinary medicines. The full mandate of BDSG can be found here. In 2023 the BDSG met eight times virtually and one time in person.

"By delivering a regulatory system able to integrate Big Data into its assessment and decision making, we can support the development of innovative treatments more quickly and optimise the safe and effective use of medicines." **Big Data Task Force Final Report, January 2020**

**Official address** Domenico Scarlattilaan 6 ● 1083 HS Amsterdam ● The Netherlands
**Address for visits and deliveries** Refer to www.ema.europa.eu/how-to-find-us
**Send us a question** Go to www.ema.europa.eu/contact **Telephone** +31 (0)88 781 6000
An agency of the European Union

# 2023 highlights for the priority recommendations

Figure 1 below provides a summary of the key BDSG highlights presented in the context of the priority recommendations.




## *Description of 2023 Highlights*

## Recommendation 1: DARWIN EU®

DARWIN EU ® coordination centre celebrated its 1st birthday in February 2023 and continued the establishment of the network throughout 2023 which realised its promise to be instrumental in enabling the use of RWE. 20 data partners either already active or being onboarded are now providing access to data from millions of patients from 12 EU Member States. The list of newly onboarded data partners will be published on the DARWIN EU website in due course.

Standard analytical pipelines and code lists (phenotypes) have been established for DARWIN EU. Industry provided useful feedback on the Catalogue of Standard Data Analyses. In agreement with the DARWIN EU Advisory Board, the processes to inform or consult industry on studies were agreed in 2022 and implemented in 2023.

After the delivery of 4 studies in 2022/2023 to support EMA's Pharmacovigilance Risk Assessment Committee (PRAC), EMA's Committee for Orphan Medicinal Products (COMP) and EMA' Committee for Medicinal Products for Human Use (CHMP), more studies have been initiated and delivered in 2023 (the second year of establishment). Currently, 18 DARWIN EU studies are ongoing or completed. Pilots with EMA scientific committees, the European Centre for Disease Prevention and Control (ECDC) and HTA bodies/Payers, and studies with the EU Vaccine Monitoring Platform were newly initiated or continued in 2023.

Study protocols and reports are published on the EU PAS Register once their assessment has been completed by EMA. Some of the most recently completed studies include:

- Drug utilisation study (DARWIN EU®) investigating treatment patterns of drugs used in adult and paediatric population with systemic lupus erythematosus.

- Drug utilisation study aiming at understanding the utilization pattern of endothelin receptor antagonists (ERAs) and phosphodiesterase-5 inhibitors (PDE-5is) in pulmonary arterial hypertension (PAH).
- Drug utilisation study investigating the annual incidence and annual period prevalence of use of opioids, in 7 European Countries, and determining duration of prescription opioid use, characteristics of new users and indication for prescribing/dispensing.
- Drug utilisation study addressing Take-Home Naloxone (THN) use to treat opioid overdose.
- Cohort study evaluating multiple myeloma patients' characterisation, treatments, and survival in the period 2012-2022.
- Cohort study estimating background rates of serious adverse events and causes of death in severe asthma.
- Natural history study of dermatomyositis (DM) and polymyositis (PM) in adults and paediatric populations.
- Drug utilisation study of medicines with prokinetic properties in children and adults diagnosed with gastroparesis

DARWIN EU continued to be supported by the DARWIN EU Advisory Board (its mandate and membership were reviewed in 2023) to provide strategic advice and recommendations on the establishment of DARWIN EU ®, its possible integration as a node within the European Health Data Space (EHDS), on how to ensure coordination and alignment with relevant European and EU Member State initiatives and policies, and how to optimise communication on DARWIN EU with stakeholders. Agendas and minutes are published on the DARWIN EU ® webpage.

In 2023, DARWIN EU took part in the EHDS2 pilot with its use case on coagulopathy and COVID-19. EMA also continued to engage with the EHDS 'Joint Action' TEHDAS to ensure alignment.

Access to and analysis of real-world data complements evidence from randomised clinical trials to enable timely and reliable evidence for the development, authorisation and supervision of medicines for patients.

## Recommendation 2: Data quality and representativeness

The first EU Data Quality Framework for EU medicines regulation has been developed with the Methodology Working Party and was published in 2023, following a public consultation.

Informed by a multi-stakeholder workshop, co-organised with TEHDAS, this framework document sets out the criteria for a more consistent and standardised approach to the quality of data used in medicine regulation and is intended to be an overall umbrella from which more focused recommendations will be derived for specific data domains.

This will help to further develop data quality assessment procedures and recommendations for current and novel data types, support pharmaceutical companies and other stakeholders in selecting data sources for their studies and support the trust of patients and healthcare professionals in data-driven regulatory decision-making.

This framework builds on the data quality recommendations of TEHDAS.

Moving forward, EMA will work with relevant stakeholders to utilise the framework's concepts and develop practical guidelines for assessing the quality of different types of data. These guidelines will

initially focus on the domains of real-world data (RWD) followed by adverse drug reaction reports (ADR).

The BDSG participated to the EMA multi-stakeholder workshop on qualification of novel methodologies that explored the scope, process and outcomes of the qualification of novel methodologies platform. A follow up Workshop on Patient Registries will be held on 12/13 February 2024.

By understanding and increasing data quality, the selection of data and interpretation of study results is informed, and the evidentiary value of studies can be judged.

## Recommendation 3: Data discoverability

In 2022 the first metadata list for real-world data sources and studies was endorsed and published. Building on the list and with the objectives to increase the discoverability of real-world data sources and increase transparency of real-world studies, the HMA EMA catalogues of real-world data sources and non-interventional studies will be rolled out to the public in the first quarter of 2024. These will be accompanied by the final Good Practice Guide for the use of the Metadata Catalogue of Real-World Data Sources and user guide.

The catalogue of data sources is set to replace the ENCePP Resources Database and provide a centralised repository for real-world data sources. At the same time, the catalogue of studies will replace the current EU PAS Register®, boosting transparency in non-interventional studies. Improvements reflected in the new catalogues will include the use of the agreed set of metadata and the link between a given data source and the study using it. Users will also benefit from a more modern technology with enhanced data submission, view, search and export functionalities. Users will be required to register to submit and manage their content in the catalogues. Through these public searchable catalogues, data will be more discoverable and of known quality and representativeness. Stakeholder engagement will intensify in 2024 to collect detailed metadata on selected data sources to populate the catalogues of Real-World Data sources and non-interventional studies. Collaboration with the joint-Action 'TEHDAS' continued in 2023 including on describing data sources in anticipation of the future European Health Data Space.

Reinforcing patient relevance in evidence generation is a key priority in the EMA's Network Strategy and the realisation of a data-driven regulatory network. BDSG contributed to initial discussions, along other stakeholders and experts, on the need for a reflection paper on the best EU approach to define, generate, collect and analyse Patient Experience Data (PED). The drafting of this reflection paper will be further progressed in 2024, including a public consultation. The BDSG also discussed how key opportunities in its workplan can be leveraged to enhance the use of PED for regulatory decision making. The HMA EMA catalogues of Real-World Data sources & non-interventional studies launching early in 2024 represent an opportunity to make PED more visible and findable. The Data Quality Framework for EU medicines regulation will include a deep dive on RWD in 2024 and presents a further opportunity through its application to PED. A research programme on establishing the value of PED is currently under discussion and will be discussed in the future with stakeholders.

Agreement on metadata to describe and identify data sets is enabling data discoverability (including through a publicly available catalogue of real-world datasets) and increasing the ability to judge the evidentiary value of non-interventional (observational) studies and real-world data sources.

## Recommendation 4: EU Network skills in Big Data

After the selection of the training providers, the Big Data training curricula have been launched in December 2023 with the rollout of the first two modules related to the Pharmacoepidemiology and Real-World Evidence (RWE) domain. Members of the EU medicines regulatory network can now access the first modules through the EU Network Training Centre Learning Management System (EU NTC LMS).

The *Introduction to Real World Evidence* module introduces the roles of RWE in regulatory practice and studies that can be conducted in this context. The *RWD Sources* module provides an overview of the different data sources that can be used to generate RWE, focusing on aspects such as data access, data heterogeneity and data quality. Overseen by the Methodology Working Party, the Big Data curricula will also include trainings in the domains of biostatistics and data science.

Further modules in the Pharmacoepidemiology/RWE curriculum will be published over the next two years. Work is also ongoing on the development of a Data Science curriculum, with the first modules in this curriculum expected to be released in January 2024.

Training is supporting the development of an expert workforce able to advise on and interpret big data.

## Recommendation 5: EU Network processes

The value of RWE is being established, use case by use case, and a solid demand for RWD studies continued to be seen in 2023.

Learning not only from DARWIN EU® studies but also from other RWE pilots, EMA published in 2023 a report on its experience in using real-world evidence to support regulatory decision-making between September 2021 and February 2023. During this period, 61 real-world data (RWD) research opportunities were identified focusing on medicine safety, medicine use, disease epidemiology, design and feasibility of clinical trials and clinical management, 30 pilot studies were initiated and 27 completed. The report establishes that RWE from studies led by regulators can complement evidence from other sources, such as randomised clinical trials. RWE can also support both pre-authorisation and post-approval assessments conducted by EMA's scientific committees and working parties. This review will be repeated in future years to monitor the progress as DARWIN EU® is further scaling up.

Overall, in 2023, 58 studies were requested or offered, of which 29 studies were considered feasible to be conducted via the Agency's in-house databases (8), framework contracts (3) or DARWIN EU® (18).

Of these, 27 studies were initiated in 2023 including 1 study for COMP, 4 studies for PDCO, 3 studies for CHMP, 8 studies for PRAC, and 1 study for the Executive Steering Group on Shortages and Safety of Medicinal Products (MSSG). Of the remaining 10 studies, most were conducted in collaboration with the extended network of decision makers including HTA/payer organisations (2), the European Centre for Disease Prevention and Control (ECDC) and the Vaccine Monitoring Platform (VMP) (6), and the European Commission (1, EHDS pilot).

Most of the studies are descriptive in nature, often investigating drug utilisation and/or medicines safety.

Several studies initiated in 2023 were linked to broader research projects that will continue in 2024:

- six studies were commissioned in accordance with the Vaccine Monitoring Platform (VMP) research agenda, looking at the effectiveness of bivalent COVID-19 vaccines (2 studies);

background incidence rates of adverse events of special interest across vaccine products; age-specific incidence rates of RSV-related disease; and effectiveness of Human Papillomavirus vaccines against cervical cancer; and a framework for the post-authorisation safety evaluation of vaccines in the EU.

- two studies were requested and are explored in line with the CHMP 2023/2024 workplan item on geriatrics (frailty/polypharmacy and inappropriate prescribing),
- a pilot study was conducted to develop a framework for RWE to support regulatory decisions for paediatric extrapolation (natural history of dermatomyositis and polymyositis),
- two HMPC pilot studies are currently being explored for their feasibility via DARWIN EU® (use of medicinal cannabis and use and safety of Pelargonium radix in children),
- two studies were requested within the mandate of MSSG to support the monitoring of prescription of medicines for public health emergencies at risk of shortages and inform on medicines used in mechanically ventilated patients.

In addition, two HMPC pilot studies were explored for their feasibility via DARWIN EU® (use of medicinal cannabis and use and safety of Pelargonium radix in children) and may lead to specific studies in 2024.

Learning from pilots is informing process improvement, guidance development and the establishment of evidentiary value. Enabling high quality and rapid assessment of medicines will improve decision making throughout the lifecycle of medicines and additionally support greater preparedness for health crisis response.

## Recommendation 6: EU Network capability to analyse Big Data

*Clinical trial raw data*

The CHMP clinical trial raw data proof-of-concept (PoC) pilot was launched in 2022 to help assess the benefits and practicalities of access to clinical trials raw data in the assessment of medicines. The PoC pilot has selected its fifth regulatory procedure and reached the halfway point in its journey to test the role of analysis of clinical trials 'raw data' submitted to EMA in electronic structured format. To date, the PoC pilot includes the following types of procedure:

- Procedure #1: initial marketing authorisation application in the area of neurology.
- Procedure #2: biosimilar application in the area of endocrinology.
- Procedure #3: variation application in the area of oncology.
- Procedure #4: type II variation application in the area of dermatology.
- Procedure #5: initial marketing authorisation application in the area of gastroenterology.

The PoC pilot is supported by the Network Advisory Group on Raw Data and the Industry Focus Group on Raw Data. Support to MAH has been important, and EMA published in 2023 the application of EMA's transparency principles to the raw data proof-of-concept pilot. The Network Community on Raw Data has also been established in 2023.

Initial insights will be gathered into an interim PoC pilot report, a summary of which is expected to be published in early 2024.

*Non-Clinical raw data*

In 2023 the EMA Proof of Concept (PoC) study to evaluate implementation of the Standard for Exchange of Nonclinical Data (SEND) was discussed with BDSG members and included in its updated workplan. The PoC will look at how SEND could improve the quality and efficiency of routine assessment but also of procedures where data is complex and requires visualization or independent analysis and how it could be useful for procedures where rapid regulatory input is needed (e.g. rolling review, accelerated procedures, PRIME) and for regulatory science projects or read-across activities to facilitate harmonization, policy or guideline recommendations.

*AI*

In 2023 the first EU AI reflection paper to support the safe and effective development, use and regulation of human and veterinary medicines has been published as draft for public consultation. The paper, co-developed by the BDSG, the Committee for Medicinal Products for Human Use (CHMP) and its Methodology Working Party, and Committee for Veterinary Medicinal Products (CVMP) outlines principles relevant to medicines for which AI would be used at any step of a medicines' lifecycle, from drug discovery to the post-authorisation setting.

BDSG discussed and learned from national and EMA initiatives on AI and in December 2023, EMA and the Heads of Medicines Agencies (HMAs) published the Artificial Intelligence workplan to 2028, setting out a collaborative and coordinated strategy to maximise the benefits of AI to stakeholders while managing the risks. The AI Workplan to 2028 covers activities across four domains (guidance, policy and product support; tools and technology; collaboration and change management; experimentation). It ensures the EMRN remains at the forefront in benefiting from AI in medicines regulation and will help the regulators to embrace the opportunities of AI for personal productivity, automating processes and systems, increasing insights into data and supporting more robust decision-making to benefit public and animal health.

The second Joint HMA/EMA AI workshop took place in November 2023 to discuss with stakeholders the latest developments in artificial intelligence (AI) technology, policy, and its potential applications in medicines regulation. Activities undertaken by HMA/EMA to address stakeholder priorities on AI, as well as the draft AI reflection paper and AI Workplan to 2028 were also discussed.

*Additional analytics initiatives*

Several initiatives to experiment on advanced analytics, including AI, have been launched under BDSG umbrella.

Scientific Explorer is an AI enabled tool for EU regulators that will facilitate easy, focused and precise searches within regulatory procedure documents submitted to regulators. It supports scientific decision-making by providing access to relevant scientific information. In 2024 the EMRN will roll-out this AI knowledge management tools for core regulatory processes, starting with Scientific Advice to support medicinal product R&D.

Individual medicines agencies of the European Network also started to collaborate to share good practice in data and analysis (in the area of data access, legal aspects, capabilities, infrastructure, methods development and Artificial Intelligence). Led by DKMA and based on expert discussions involving multiple National Competent Authorities, the discussion paper on Clusters of Excellence has been endorsed by the BDSG and published in March 2023.

BDSG also provided a place to discuss and learn from the European Commission research activities in the field of Big Data and RWE, including the 5 projects (Real4Reg, REDDIE, More Europa, ONCOVALUE and REALM projects) from the HORIZON-HLTH-2022-TOOL-11-02 call on 'New methods for the

effective use of real-world data and/or synthetic data in regulatory decision-making and/or in health technology assessment'.

Demonstrating value of raw data analysis and fostering knowledge and expertise within EU network is key to enable high quality and rapid assessment of medicines.

## Recommendation 7: Delivery of expert advice

The Methodology Working Party (MWP) was established in 2022 and established its 3-year workplan (2nd version was opened for public consultation in Q4 2023). The methodology European Specialist Expert Community (ESEC) has been established with more than 180 experts. The ESEC will strengthen the MWP, bringing together a broad range of expertise including Big Data, biostatistics, real world evidence, advanced analytics (including AI), PK/PD, extrapolation, modelling and simulation, GCP and omics.

To ensure high quality decision making in this rapidly developing environment, the MWP is currently drafting a reflection paper on the use of RWD to generate RWE in non-interventional studies and is establishing a roadmap for the development of RWE guidance.

Expert advice including on advanced analytics, real world evidence and 'omics is empowering robust assessment and decision-making by regulatory committees.

## Recommendation 8: Governance framework

In 2023 the review of the Network data governance was completed. HMA and EMA Management Board endorsed revised mandates of the two main data governance bodies of the EMRN - the Network Data Board and the Big Data Steering Group. The revisions reflect evolution to the mandates of these bodies to clarify their scope, division of responsibilities, composition and stakeholder engagement models. The BDSG mandate and membership is now strengthened with representation of EHDS, CTCG, HTA/Payers, ethic bodies or networks (nominations process is now also completed) and an explicit mandate for AI.

The BDSG continued to prepare for a changing policy environment with the future European Health Data Space (EHDS) and the revised Pharmaceutical Strategy for Europe via regular updates from the European Commission and participation to various fora and workshops.

The BDSG has annually updated its workplan (BDSG workplan 2023-2025), informed by stakeholders and partners feedback.

A series of five training webinars on Data Protection in medicines and public health has been delivered by EMA to experts from Members States. These are also published on the EU Network Training Centre.

Ethics considerations have been included in the AI reflection paper launched for public consultation in 2023.

Secure and ethical data governance is an enabler for secondary use of healthcare data and the work of the BDSG seeks to support stakeholders to navigate and comply.

## Recommendation 9: International initiatives

Progress in convergence with international partners on RWD/RWE continued in 2023 and contributes to establish the value of RWE and enabling its use. Numerous initiatives are underway:

- Building on the 2022 ICMRA workshop hosted by EMA, in June 2023 the ICH Assembly endorsed the public consultation for a new Reflection Paper co-led by EMA, FDA and Health Canada, on "International Harmonization of RWE Terminology and Convergence of General Principles Regarding Planning and Reporting of Studies Using RWD, with a Focus on Effectiveness of Medicines". It proposes two ICH guidelines: the first to harmonise RWD/RWE terminology, metadata and general principles for regulatory assessments, and the second, to harmonise the principles for structure and content of protocols and reports and recommended "best practices" for registration of study protocols/results.
- The public consultation on the draft ICH M14 Guideline on "General Principles on Planning and Designing Pharmacoepidemiological Studies That Utilize Real-World Data for Safety Assessment of a Medicine" is expected in Q1 2024.
- The existing ICMRA COVID-19 RWE Working Group has been re-purposed to focus on RWE in public health emergencies.

Convergence with international partners on standards and guidelines will leverage best expertise and minimise burden on stakeholders.

## Recommendation 10: Stakeholder engagement

Collaborations with external stakeholders and partners continued in 2023.

The multi-stakeholder workshop on qualification of novel methodologies was held in April 2023 to explore the scope, process and outcomes of the qualification of novel methodologies platforms. It will inform further discussion with BDSG.

The workshop on Real World Data (RWD) quality and Real World Evidence (RWE) use was held in June 2023 to address RWD quality linked to the EU Data Quality Framework as well as recent use of RWE in the regulatory decision making process.

The fourth annual HMA/EMA Big Data Stakeholder Forum 2023 was held in December 2023 to update participants on the delivery of big data activities under the Network Strategy to 2025. Discussions with stakeholders took place on further opportunities for big data in medicines development and regulatory decision-making. In addition, keynote speakers shared their insights on cutting-edge topics.

The Methodology Working Party stakeholder interaction meeting was held in December 2023 to consolidate comments received during the public consultation on the revised Methodology Working Party (MWP) workplan. The objective of the meeting was to shape the updated MWP workplan for final endorsement at the Committee for Medicinal Products for Human Use (CHMP) preparatory and organisational matters (PROM) plenary meeting.

Two dedicated BDSG industry stakeholder meetings (in May and November 2023) were organised in the context of EMA's continuous endeavours to foster regular interactions with industry stakeholders on topics of common interest. Agenda and minutes are published on EMA website.

The EMA Big data newsletter, published 4 times in 2023, has provided an update on progress in implementing the workplan of the HMA-EMA Big Data Steering Group.

Listening to stakeholders and leveraging their work is optimising and maximising transformation to data-driven regulation. Transparency and communication build trust in what is delivered.

## Recommendation 11: Veterinary recommendations

Following the adoption of the European Veterinary Big Data strategy 2022- 2027 in 2022, the EU Veterinary Big Data Workplan to 2022-2025 transposes the strategy's pillars into the actionable workstreams (analytics discoverability, governance and literacy, stakeholder engagement) with the overall ambition of prioritising digital veterinary use cases for the coming years at EU level. The workplan was adopted and published in 2023.

Established in June 2023, the Veterinary Data Hub brings together a multidisciplinary team of Network experts from Belgium, France, Germany, Portugal, Spain and Sweden. This group of experts collaboratively engages in advancing the EU Veterinary Big Data strategy and contributes to the planned activities within this framework.

The Third Veterinary Big Data Stakeholder Forum was held in November 2023 to continue the discussion on Big Data in veterinary medicines regulation. The event reported on the EU Veterinary Data Hub activities and the HMA/EMA Veterinary Big Data workplan, discussed use cases, future needs and upcoming challenges of big data in the veterinary regulatory domain.

In the veterinary domain, a research study, 'Big Data in Veterinary Medicines Regulation: A Data Landscape Analysis', was launched in 2023 to map the data landscape and identify relevant data sources. This is planned to be followed by a study on metadata and subsequently the information on veterinary data sources will also be included in the catalogue of data sources.

Synergies exist in the use of data between the human and veterinary domains that can catalyse our transformation.