09 January 2023
EMA/826814/2022

Big Data Steering Group (BDSG): 2022 report

This report provides a summary of the key activities and achievements of the BDSG in 2022. Significant progress in the transformation to data-driven regulation continued in 2022, in line with the Network Strategy to 2025 and BDSG workplan.

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.

This report is based on the second BDSG workplan. In 2022 the BDSG met 10 times virtually.

Step by step, progress is made in line with the agreed workplan on enabling the use and establishing the value of the Big Data in the development, authorisation and supervision of medicines in Europe.

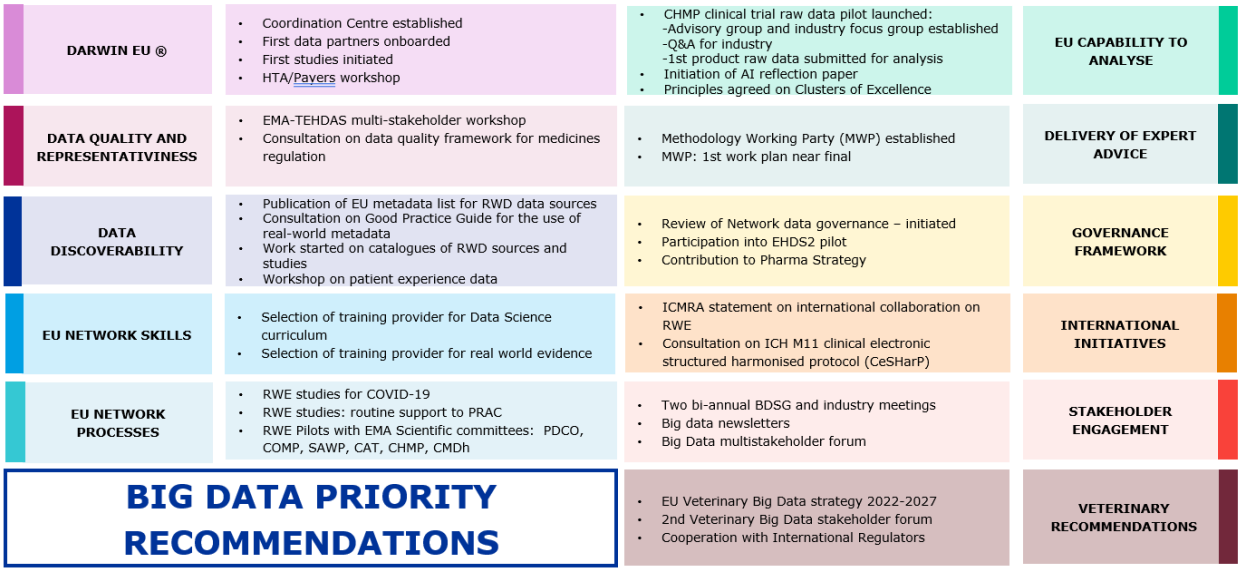
Informed by feedback from experts and stakeholders, the third workplan covering 2023 to 2025 was adopted by the BDSG in 2022.

"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

2022 highlights for the priority recommendations

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



Description of 2022 Highlights

Recommendation 1: deliver a sustainable platform to access and analyse healthcare data from across the EU (Data Analysis and Real-World Interrogation Network – DARWIN EU®)

DARWIN EU ® is now operational with the appointment and establishment of a consortium led by Erasmus University Medical Center Rotterdam as the DARWIN EU ® coordination centre.

The first data partners have now been onboarded and provide access to EMA and the EU regulatory network to more than 26M patients' data from sources such as hospitals, primary care, health insurance, registries and biobanks. DARWIN EU ® is expected to bring on board a minimum of ten additional data partners every year.

As planned, the first 4 studies have also been initiated by DARWIN EU® to support EMA scientific committees (COMP, PRAC, CHMP): an epidemiological study of rare blood cancers to inform on their prevalence in Europe; a drug utilization study of valproate-containing medicinal products; a study on the use of antibiotics to inform future work on anti-microbial resistance; and a study on all-cause mortality rates in patients with severe asthma.

More regulatory use cases were explored this year with HTA bodies and payers as well as the European Centre for Disease Prevention and Control (ECDC) to identify study proposals to be piloted via DARWIN EU®.

As part of the establishment of DARWIN EU®, a data protection impact assessment was performed and lessons learned were shared with experts from national competent authorities on how data protection matters were addressed during the procurement and establishment phases of DARWIN EU® in accordance with Regulation (EU) 2018/1725, the European Union Data Protection Regulation (EU DPR) and Regulation (EU) 2016/679, the General Data Protection Regulation (GDPR).

The DARWIN EU ® Advisory Board met regularly in 2022 to provide strategic advice and recommendations on establishing DARWIN EU ® and its future use of 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 ®. Agenda and minutes are published on the DARWIN EU ® webpage.

As DARWIN EU ® will act as a pathfinder initiative and use case for the proposed EHDS, EMA is actively collaborating with the European Commission. In 2022, an EMA use case was selected to participate in the EHDS2 pilot to demonstrate the potential of transnational reuse of health data for research, innovation, development of regulations and policies and, ultimately, personalised healthcare. 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: Establish an EU framework for data quality and representativeness

Informed by a multi-stakeholder workshop, co-organised with TEHDAS, the first draft EU Data Quality Framework for medicine regulation was launched for public consultation in 2022.

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, e.g. real world data in 2023.

This will help to help 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 ensure the trust of patients and healthcare professionals in data-driven regulatory decision-making.

This framework builds on the data quality recommendations of TEHDAS.

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: Enable data discoverability

In 2022, the first metadata list for real-world data sources and studies was published to improve transparency with regard to the discoverability of data sources and real-world studies.

The metadata for data source covers administrative details, the data elements collected, quantitative descriptors, information on data flows and management, as well as vocabularies. The metadata for studies covers administrative details, information on methodological aspects and data management.

The metadata will be included in two catalogues that are under construction: the catalogues of observational data sources containing information about existing real-world databases (to replace the current ENCePP catalogue); and the catalogues of observational studies (to replace the current EU PAS Register). Work is progressing to establish these two catalogues with the launched expected in late 2023 or early 2024.

Work is ongoing to collected detailed metadata on selected data sources to populate catalogues of observational data sources.

To prepare stakeholders in anticipation of the go-live of the catalogues, a Good Practice Guide for the use of real-world metadata was released for public consultation in 2022. It aims to help regulators, data holders, researchers, pharmaceutical companies and other interested stakeholders to use the future catalogue of data sources and provides recommendations on how to identify suitable real-world data sources for studies.

Alignment with TEHDAS recommendations on describing data continued in 2022.

BDSG also learned from the workshop on Patient experience data in medicines development and regulatory decision-making .

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 observational studies and real-world data sources.

Recommendation 4: Develop EU Network skills in Big Data

Following adoption of the Big data training curricula last year, the selection (competitive tender) for training providers to deliver the trainings was launched in 2022. A training provider on pharmacoepidemiology and real world evidence has been selected and the selection of a provider on data science is being finalised.

Trainings in line with the curriculum will be rolled out to the EU Network via the EU Network Training Centre to increase the breadth and depth of EU Network knowledge and experience in data science, methods and analytics to advise companies developing products and to expertly assess application dossiers.

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

Recommendation 5: Strengthen EU Network processes for Big Data submissions

The pilots of generating RWE for use cases through the product life cycle have progressed with the EMA scientific committees (PDCO, COMP, SAWP, CAT, CHMP, CMDh) and will continue into 2023.

Overall, in 2022, 50 studies have been requested and/or offered of which 25 studies were feasible to be conducted via the Agency's in-house databases, framework contracts (natural course of spinal muscular atrophy and its standards of care overtime) or DARWIN EU® (noting that the onboarding of first data partners in DARWIN EU® only took place towards the end of 2023). The proportion of feasible studies will increase steadily as additional data partners are onboarded by DARWIN EU®.

Several studies were completed in 2022 (PDCO (2), COMP (3), SAWP (2), Other/repurposing pilot (1)) while others are still ongoing (CAT (1), CHMP (1), COMP (1), PDCO (1), NCA (1)). Two of these studies are performed via DARWIN EU®, including an analysis to investigate the prevalence of rare blood cancers (COMP) and another aiming at determining background rates of mortality and serious cardiovascular events in patients with severe asthma (CHMP). Most of these analyses were descriptive disease epidemiology studies looking at incidence or prevalence rates as well as natural history of disease and treatment patterns. Other studies intend to inform drug utilisation and/or the design and feasibility of clinical trials.

Standard processes to routinely support PRAC with RWE studies have now been implemented and 12 studies have been completed in 2022.

In the context of COVID-19, RWE studies have been contributing to the collective evidence on the benefit-risk of vaccines, via EMA framework contracts with large research organisations/consortia, complementing the monitoring by vaccine developers and by Member States. 11 RWE studies have been set up between 2019-2022, of which 6 are completed and 5 are ongoing. Additional studies to address further COVID-19 research questions are being planned.

Table 1: COVID-19 EMA-funded RWE studies (2019-2022) as of December 2022

Studies	Status	EU PAS Register no. Publications (as of Dec. 2022)
Readiness EU infrastructure for COVID-19 vaccine monitoring ('ACCESS') • Background incidence rates of AESIs • Template protocols for vaccine safety and effectiveness studies • Feasibility of monitoring vaccine coverage, safety and effectiveness in EU healthcare databases	Completed	EUPAS37273 EUPAS39370 EUPAS39361 EUPAS39289 Willame et al. 2022
Multicentre collaboration for COVID-19 patient medication cohort studies ('E-CORE')	Completed	EUPAS38759
Readiness Impact of COVID-19 infection and medicines in pregnancy ('CONSIGN').	Ongoing	• WP1 (EHRs): EUPAS39438 • WP2 (COVI-PREG): EUPAS39226 • WP3 (INOSS): EUPAS40489 • Meta-analysis: EUPAS40317
Natural history of coagulopathy and use of anti-thrombotic agents in COVID-19 patients.	Completed	EUPAS40414 Burn et al. 2022 (1) Burn et al. 2022 (2)
Early safety monitoring (*Early-Covid-Vaccine-Monitor*/'ECVM') • Prospective in vaccinees (WP1): BE, SK, FR, DE, IT, NL, UK • EHRs (WP2): databases in ES, IT, NL, UK	• WP1: Ongoing • WP2: Completed	WP1: EUPAS39798 WP2: EUPAS40404 Sturkenboom et al. 2022 (medRxiv)

Extended safety monitoring (*Covid-Vaccine-Monitor*/'CVM')
• Prospective in vaccinees:
  ○ WP1 (special populations): NL, IT, PT, RO, SK, ES, CH, HR
  ○ WP2 (general population): NL, DE, BE, FR, IT, HR, RO, SK, IE, CH, ES
• EHRs (WP3/WP4): framework for signal strengthening and methodology, 9 databases in IT (3), ES (3), NL (1), UK (1), NO (1)
	Ongoing	EUPAS42504 (WP1)
		EUPAS39798 (WP2)
		EUPAS42467 (WP3/WP4)
		Bots et al. 2022

Association between thromboembolic events and COVID-19 vaccines	Completed	EUPAS44469 Li et al. 2022 Xie et al. 2022

Studies	Status	EU PAS Register no. Publications (as of Dec. 2022)
Comparative effectiveness of heterologous and homologous primary- and booster SARS-CoV-2 vaccination schedules in the Nordic countries	Completed	EUPAS46537 Andersson et al. (medRxiv)
Effectiveness of COVID-19 vaccination in 5 EU countries	Ongoing	EUPAS47725
Association between COVID-19 vaccines and paediatric safety outcomes in children and adolescents aged 5-19 in the Nordic countries	Ongoing	EUPAS48979
Impact of EU label changes and regulatory communication on SARS-CoV-2 adenovirus vector vaccines in context of thrombosis with thrombocytopenia syndrome (TTS): risk awareness and adherence	Ongoing	EUPAS44970

The number of requests for studies and studies delivered attest of the solid demand for RWE studies and an increasing demonstration of their value to support EU regulatory decision-making process.

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: Build EU Network capability to analyse Big Data

Clinical trial raw data

Having established the necessary processes and conducted a data protection impact assessment, the CHMP raw data pilot was launched in 2022. The pilot is exploring the business case and practicalities for analysis clinical trials raw data in Marketing Authorisation Applications (MAAs) and is expected to last up to two years, and include 10 regulatory procedures submitted to EMA, with primary focus on initial MAAs submitted to EMA.

Following a procurement procedure, the Danish Medicines Agency (DKMA) was selected to support the pilot's execution and to conduct individual patient data analyses for three out of the ten regulatory procedures included in the pilot. The support to the other pilot procedures will be undertaken by either NCA assessment teams or EMA staff, to explore the different operational models to analyse raw data.

The first product data for an initial marketing authorisation application for the treatment of a neurological disorder was received by EMA in October 2022.

Engagement with stakeholders included the establishment of a multi-disciplinary Advisory Group on raw data with members from CHMP, EMA Working Parties, patients' representatives, as well as an Industry Focus Group. A Questions and Answers about the raw data proof-of concept pilot for industry was published.

AI

Following the 2021 Joint HMA/EMA workshop on artificial intelligence (AI) in medicines regulation, Network experts have started to draft an AI reflection paper. It will cover regulatory aspects of AI/ML including:

- Impact on assessing the Benefit/Risk in medicine development and authorization in relation to GxP standards
- Impact on post-marketing use, pharmacovigilance surveillance and drug effectiveness studies
- Model development and deployment
- Risk exploration
- Data privacy, legal and ethical recommendations
- Regulatory interactions

Additional analytics initiatives

As a complement to the work on raw data, a pilot has also been completed to leverage unstructured information in documents submitted to regulators as part of scientific advice requests and MAAs and to enable easier and more accurate identification of patterns across procedures.

The enhancement of safety data analysis to support decision making in the Eudravigilance system has been launched in 2022.

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, the discussion paper on Clusters of Excellence has been endorsed by the BDSG and drew on expert discussions including DKMA, BfArM, PEI, AEMPS, Infarmed, SMPA and MEB.

BDSG also provided a place to discuss with the European Commission research activities in the field of Big Data and RWE, including the projects to deliver HORIZON-HLTH-2022-TOOL-11-02: 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: Modernise the delivery of expert advice

In 2022, the Methodology Working Party (MWP) was established. The MWP members have been appointed and since April 2022 virtual meetings are held twice monthly. Its first workplan (covering a three-year period) is now finalized.

A methodology European Specialist Expert Community (ESEC) is in the process of being established with the participation of the various Network experts. The ESEC aims to strengthen the MWP bringing together a broad range of expertise including Big Data, biostatistics, real world evidence, advanced analytics, PK/PD, extrapolation, modelling and simulation, GCP and omics.

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

Recommendation 8: Ensure data are managed and analysed within a secure and ethical governance framework

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 also participated in a new review of the EU Network data governance which includes review of the mandates of the BDSG and EU Network Data Board. Data governance will consider the needs of stakeholders, new EU policies, the full data cycle including data sharing, interoperability and full compliance with data protection, to efficiently integrate Big Data analysis and novel technologies in the EU regulatory decision-making process.

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: Collaborate with international initiatives on Big Data

International collaboration on RWE has continued in 2022, including via the International Coalition of Medicines Regulatory Authorities (ICMRA). In June 2022 an ICMRA workshop was held for regulators to share experience in obtaining and using real-world evidence for the assessment of medicines. As an output from the workshop, ICMRA issued a statement which highlighted the opportunity for international collaboration in four areas i.e. harmonisation of terminologies for RWD and RWE, regulatory convergence on RWD and RWE guidance and best practice, readiness to address public health challenges and emerging health threats, and transparency (ICMRA statement on international collaboration on RWE in regulatory decision making).

Additionally, a structured clinical trial protocol is also being developed by ICH M11 (ICH M11 clinical electronic structured harmonised protocol - CeSHarP). Public consultation was launched in 2022 to progress its development.

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

Recommendation 10: Create an EU Big Data 'stakeholder implementation forum'

The third EMA/HMA Big Data Stakeholder Forum was held virtually in December 2022 and continues to be the principle opportunity for stakeholders to provide feedback regarding the Big Data priority recommendations and workplan and to their perspectives on implementation. This forum is also an opportunity to discuss the proposed EHDS and research initiatives on Big Data and real-world evidence under Horizon Europe.

Two dedicated Big Data Steering Group industry stakeholder meetings were organised in the context of EMA's continuous endeavours to foster regular interactions with industry stakeholders on topics of common interest.

The EMA Big data newsletter, published every three months, 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

The European Veterinary Big Data strategy 2022- 2027 was published in 2022. It provides the strategic vision to address the challenges and opportunities that digital transformation offers in the Veterinary regulatory domain and continue building upon the overall goals set by the Veterinary Medicinal Product Regulation. It is built around five pillars: Regulatory business use cases, data literacy, stakeholders, a data landscape framework, a sustainable data ecosystem and governance framework.

The 2nd Veterinary Big Data stakeholder forum was held in November 2022 to continue the discussion on Big Data in veterinary medicines regulation by presenting progress made, present concrete use cases and identify a prioritised list of initiatives that can inform the work plan on the implementation of the Big Data strategy in the veterinary domain for the coming years.

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