17 February 2025
EMA/591124/2024
Data Analytics and Methods Task Force

HMA/EMA Big Data Stakeholder Forum 2024

Report of the HMA/EMA Big Data Stakeholder Forum - 28 November 2024

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

Table of contents

Introduction and Opening remarks

As the journey towards data-driven medicines regulation continues, a new phase for the regulatory network is starting. The value of real-world data for evidence generation is becoming more established and technological advances, e.g. in the area of artificial intelligence (AI), have occurred more rapidly than originally anticipated. At the same time, the EU Network is accelerating its efforts to strengthen data governance, to share more data across the Network and to foster high level of data interoperability and standardization.

The fifth annual Big Data multi-stakeholder forum took place on 28 November 2024 at the EMA building in Amsterdam. It aimed to inspire with keynotes speakers, discuss progress and inform the an update of work planning in the context of the Network Strategy to 2028.

The forum was opened by Prof. Rui Santos Ivo (HMA lead, Head of Agency, Infarmed), Marco Greco (President of EPF) and Lorena Boix Alonso (Deputy-Director General, DG SANTE, European Commission).

Two hundred and forty-three stakeholders and partners participated in the Forum and many hundreds followed the broadcast of the event online.

This report provides a high-level summary of the forum presentations as well as highlights from the comments and discussion with stakeholders.

Session 1: Implementation of the HMA/EMA Big Data Task Force priority recommendations

This year marked the fifth year on the journey to realise HMA EMA vision of a data-driven regulatory system. Guided by the Big data steering group workplan 2023-2025, this session looked at the key progress on the delivery of the HMA EMA big data priority recommendations and heard from stakeholders on gaps and priorities.

An overview of 2024 achievements were presented for the Human and Veterinary domains. Significant progress continued in 2024 to enable the data transformation of the European Medicines Regulatory Network (EMRN) and deliver the data activities of the European medicines agencies network strategy to 2025.

More generally the last five years BDSG achievements were celebrated:

- The EMRN can now rely on Real-World Evidence (RWE) to complement data from Clinical Trials (CT) to support decision-making across a spectrum of regulatory use cases, with HTAs/Payers/ECDC also benefiting from additional evidence. Evidence generation is planned and guided by data, knowledge, and expertise, and research question drives the choice of evidence. The DARWIN EU network increases access to patient data (> 130 millions) and conduct of RWE studies. Real-World Data (RWD) and studies catalogues makes data sources more discoverable and increases transparency on observational studies. Quality of data is known and guided by the EU data quality framework.
- The EMRN is now ready to explore the use and pilot of novel data type such as clinical trial study data, patient-experience data, social media and genomic data.
- The EMRN has now increased knowledge and experience in data (data science, RWE, Biostatistics) and methods (Methodology Working Party, ESEC, SIAs).

- The EMRN use experimentation to leverage AI. It has a collaborative and coordinated strategy to maximise the benefits of AI while managing the risks.
- EU guidance is available, aligning closely with international guidance and standards to help industry and regulators develop and supervise medicines.
- Continued compliance with data protection and ethics of data sharing, high levels of transparency, and building collaborations have enabled trust from stakeholders and partners.

Discussions with stakeholders' representatives included the learnings and future priorities from regulators, pharmaceutical industry, patients, healthcare professionals (HCP), academia and heath technology assessment (HTA) bodies. Across stakeholders, the progress since the start of the BDSG five year ago was acknowledged.

For regulators, BDSG has been very successful, and progress is being monitored so that key achievements can be used for their decision.

For Industry, future priorities include further guidance on RWE, on practical considerations in tools, on how patients' input and data are to be considered and used in regulatory decision-making. Sharing learnings and collaboration through dedicated channels are also essential.

For patients, the use of RWD is important to enable evidence-based decision and promote the development of medicines that truly met patient needs. It is also very relevant for the coordination of RWE with HTAs as reimbursement decision should be informed by real-world use to benefit patients. Community led patient's registry is another valuable source of data. Patients have a unique expertise on their condition and needs, and strong collaboration should continue. Trust and transparency are another priority. The clinical study raw data pilot is one example where better understanding of the data and opportunity for independent analysis and critical evaluation of the sponsors' interpretation may strengthen public trust and transparency. In an era of misinformation and distrust in science and governments, the need of transparency in methodologies to integrate new data sources into submission and regulatory decision-making has never been higher. Regulators must keep abreast with ethical questions that will arise as new digital models are used. Relying on data that does not include a marginalised community, minorities, people of different gender or ethnicity, will introduce significant biases and lead to inequitable health outcome and may jeopardize trust. Integrating privacy consideration and data protection by design is key. Finaly, it is also critical to continue to invest in communication and strengthening digital literacy of patients.

For HCPs, while EU has ambitious plan, it is also needed to remain pragmatic on what can be achieved. EU workforce and stakeholders need to be empowered through training, highlighting the opportunities of the EU Network Training Center (EU NTC). Progress on Big data should aim to eliminate inequalities.

For academia, sustaining and successful academic collaborations are key for a data driven regulation and future work should be linked with the European Platform for Regulatory Science Research and the EMANS to 2028.

For HTA bodies, priorities for the next years should be to support the use of existing data sources, in particular the development of disease registries and network of registries (e.g. ERNs) and of common minimal datasets for regulatory purpose. With the opportunity of the entry in force of the HTA regulation in January 2025, collaboration between regulators and HTA bodies on post-launch evidence generation (PLEG) studies should be reinforced.

Session 2: Evidence generation to advance regulatory excellence, here and now

This session explored the critical role of evidence generation in enhancing regulatory decision-making and the key progress made for the use of real-world data and clinical trial raw data. In this context, convergence with international partners also remained essential.

An overview of the totality of evidence generation introduced the EMRN shared vision for clinical evidence generation by 2030.

This session then explored in more details analysis of RWD to support regulatory decision-making. It looked at the current use cases, the needs from national competent authorities, the benefits and examples of limitations of RWD. This was illustrated by several example of DARWIN EU studies. For Paediatric systemic lupus erythematosus, how young patients with SLE are currently treated as compared to adult patients has been explored. Public Health Emergencies was another example, with collaboration between EMA and the European Centre for Disease Prevention and Control (ECDC) through the Vaccine Monitoring Platform (VMP) to coordinate vaccine use, safety and effectiveness studies. A last example discussed value of RWD for drug shortages monitoring and predictions for antibiotics. After 2 years of operation of DARWIN EU, it supports now all phase of the medicinal product authorization life cycle and showing potential for the future to support use cases for the supply medicines.

To address challenges in the RWD/RWE, and clinical trial spaces, international regulators have been collaborating on key initiatives to enhance harmonization, improve methodologies, and strengthen evidence generation. Progress continued under the umbrella of the International Coalition of Medicines Regulatory Authorities (ICMRA), building on the experience from Covid and establishing a working group on RWE for Public Health Emergencies. Under the umbrella of the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH), drafting of the M11 and M14 guidelines, of the RWE Reflection Paper and the RWE New Topic Proposal also progressed.

The Clinical study data pilot interim results and its learnings were presented. The pilot aims to determine the benefits of early clinical study data access and analysis to support the scientific assessment of medicinal products. Nine regulatory procedures have been to date included in the pilot which demonstrated added value for assessment and decision-making and has been extended to continue to gather insights and learnings. The next phase of the pilot will also be looking at data submission requirements, guidance and training, and use of open-source softwares.

Discussions with stakeholders' representatives included the views and future priorities from regulators, pharmaceutical industry, patients, academia and HTA bodies.

Industry presented its approach on using multilingual and open-sourced statistical softwares to analyse data. Collaboration within the pharma industry is increasing in that area and the first regulatory submissions have been approved by international regulators.

For patients, raw data are crucial to provide significant additional information to analyse clinical data. With the use of AI, these datasets can be assessed and analysed to identify patterns and gain a deeper understanding, especially in smaller patient groups or rare conditions that may not be visible in aggregated clinical study data. The availability of raw data may also save time in the regulatory process without compromising safety. RWD are considered fundamental, especially in the case of rare diseases or newly developed therapeutic options, and to bridge the gap between clinical trials and real-life implementation. However, there are still an imbalance in patient's access to registries and concerns about data sharing and disclosure.

For regulators, regulatory science and collaboration with academia have been a key enabler for the use of RWD. The importance of patient registries, and in particular prospective data collection, was also highlighted.

For Academia, with more access to diverse data sources, there is a need to balance standardisation with local expertise, especially important in case of multinational and collaboration studies. In the context of CT, there are also many opportunities for utilizing the potential of RWD, with more pragmatic approaches required. Better integration of clinical research and clinical practice is needed, e.g. use electronic medical records / RWD routinely for data collection in CTs. This should be supported by sandbox environment and funding for proof-of-concept studies to advance the field.

For HTAs, while the use of RWD is not new, there is a change in expectations associated with RWE, as there is an increasing reliance on accelerated clinical development strategies for innovative drug and amplification of uncertainties regarding a product's effectiveness at the time of registration. The generation of real-life data should be considered together with the clinical development plan and the use existing data sources should be reinforced. Comparative studies with a complex methodology can maximise the use of the results to accelerate development strategies, and representativity of the results in terms of patients' and prescribers' characteristics should be documented.

Session 3: Evidence generation to advance regulatory excellence, preparing for tomorrow

With a changing regulatory and technological landscape, the regulatory network needs to seize every opportunity to generate of evidence from diverse types of data. This session continued to explore the future of evidence generation in enhancing regulatory decision-making.

Recent progress on the use of pharmacogenomic (PGx) data have been discussed, presenting the outcome of the recent multistakehodler workshop organized in September 2024 and the PGx Specialised Interest Area (SIA), the pharmacogenetics expert group within the Methodology working party. Looking back at the history and upcoming regulatory steps in the field of PGx, example of the use of PGx information in the SmPCs was discussed as well as some of the challenges, e.g. in regards with the lack of instruments for genetic charaterisation. More broadly, the opportunities of using Proteomics and Metabolomics in personalized therapy were presented. Future priorities for regulatory needs should include global harmonization (of guidelines and SmPC information), education of healthcare professionals (to promote the integration of PGx into routine healthcare), consideration of polymorphisms specific to different ethnicities, conduct of more cost-effectiveness studies and leverage real-world data to strengthen PGx evidence. , PGx information should be sufficiently detailed in SmPCs and applicable in clinical practice, and further in-depth evaluation of areas such as e.g., psychopharmacology and anticoagulative therapy should be prioritised.

mHealth data is another source of data that can support evidence generation in the future. After a short description on the key characteristics of mHealth data, together with the challenges and opportunities, the potential use cases for regulatory decision-making have been discussed. Like RWE in general, mHealth data can support the planning and validity of applicant studies, understand clinical context, and investigate associations and impact. A set of points for considerations to enable further their use has been presented. Despite challenges in data quality, compliance with data protection requirements, access, potential bias, there are opportunities, in particular as mHealth tools can generate large amount of detailed patient data. mHealth data may have utility for regulatory decision-making and the EMRN will continue to monitor progress in this area, as technological advancements may help overcome some of the challenges in using such data.

Finally, the potential value of social media data for regulatory decision-making was presented. Their potential use cases are similar to the ones of mHealth data with the addition of potential uses cases to monitor, prepare for, and address public health challenges (e.g. supply and availability of Medicines, public health emergency detection, and misinformation and stakeholder communication). A set of points for considerations to enable further their use has been presented. Social media could have different utility and impact for regulatory decision-making by providing a timely monitoring tool and near real-time stakeholders' listening, by allowing to communicate directly with patients and combat misinformation. Social media data could be a complementary source of information to data-driven regulatory decision-making and act as a trigger to more detailed investigation via other regulatory pathways. Despite significant challenges (data quality, ethical considerations, differences of social media platforms, compliance with data protection requirements), some possible limitations might be overcome in the future through technological advances. Researchers in this field are encouraged to conduct further studies, as this could help enhance the EU regulatory system and bring about public health benefits

Discussions with stakeholders' representatives included the views and future priorities from pharmaceutical industry, patients, regulators and ethic.

Industry highlighted the opportunities of genomics data for precision medicine and to support drug discovery and development. There is a need for fair and simple terms of access by industry researchers, as well as agreement on how to allow for linkage while maintaining data privacy (e.g. GDPR). Data standards should be developed and published for digital tools and mHealth data, and clarity is required on evidence generation requirements, specifically when used in clinical trials. The potential value of social media should consider noises and biases and the acceptability and relative value of this type of supporting evidence in different scenarios should be clarified.

For patients, the digital divide of the real world should be considered, in particular regarding access to digital resources, that course worsen existing inequalities and introduce biases. Diverse data sources do not always equate to data diversity. To prevent this, patients should be involved in designing RWD collection methodologies, identifying meaningful RW database endpoints, and developing strategies to enhance trust and trustworthiness of data platforms and processes. Patients' organizations have a key role in public private partnership to bring patients experience in the discussions. Work on accessible communication and training is also essential.

For regulators, social media data may hold potentially valuable information in several contexts but their use in medicines evaluation is not standard practice yet in EU medicines regulation and remains very limited. There is a need for focused initiatives and proof of concept studies to assess social media data utility across various regulatory applications (e.g. safety and misuse). The role of digital, specifically mobile devices in drug development and regulation, is only expected to grow but not without challenges. To overcome some of these challenges, technical expertise requirements/capacity building, early dialogue with regulators and guidance are needed. For genomic data, efforts should continue to share learnings and build knowledge amongst regulators, to expand the network's access to more genomic data. The EHDS should be leveraged to harness the increasing amount of pharmacogenomic information and integrate it with other sources of data, such as electronic health records.

Ethics should be integrated at every step of the regulatory network work, from research to policy to implementation. Ethics should be considered by design for tools. Transparency, openness and documentation of points of views/positions are essential.

Session 4: Unlocking the value of data with the EMRN network data strategy

Increasing the value of data for the benefit of public and animal health is the network vision on data. This session explored how to maximise the generation of evidence, the use and exchange of data to support regulatory decision-making. For this, a robust network data strategy is essential to ensure the network data assets are appropriately managed and have a high level of interoperability, standardisation and quality.

For the EU network to fully realise the value from data to benefit public/animal health and its business processes, the network data asset needs to be managed appropriately. The ERMN data strategy currently under publication consultation was introduced to stakeholders. It aims to provide the vision, principles and objectives for implementing effective data management so that the EMRN can be transformed into a data-driven regulatory ecosystem that promotes innovation, efficiency, and better coordinated decision-making among national authorities and the EMA. They key strategic objectives of the strategy are related to data governance, data quality management, interoperability, data cataloguing and metadata management, knowledge and change management, and analysis of data. The two existing data governance groups, the BDSG and the Network Data Board, have been unified into one network data governance group, the Network Data Steering Group (NDSG) that will start in 2025. The NDSG will also support the delivery of the EMAN strategy to 2028 (including Theme 2: Leverage data, digitalization and AI).

Each objective of the Network data strategy was then discussed in more details, illustrated by examples of ongoing initiatives and national or EU level, e.g. the HMA-EMA Real-World Data Catalogues, the European Health Data Space data catalogues, the existing EU Data Quality Framework and UPD Data quality Framework, the European Interoperability Framework and Interoperable Europe Act, progress on SPOR Master data services and ISO IDMP standards, the HMA/EMA data standardisation strategy, and experience from the Norwegian Medical Products Agency in implementing advanced AI analytics tools.

Discussions with stakeholders' representatives included the views and future priorities from pharmaceutical industry, patients, European Commission (EC), standard development organizations (SDOs) and ethics.

For pharmaceutical Industry, the ERMN data strategy is welcome and help bringing together many data experts, bridging the gaps between different functional silos. Priorities for the future should focus on data quality, transparency on data and studies, and international harmonisation. Area for collaboration and contribution include accessibility of data for all stakeholders and pilot to analyse data.

For Patients, the EMRN data strategy should be an opportunity for simplification in this complex ecosystem, to reduce fragmentation and duplication of effort, and to increase the coordination of the different stakeholders. Patients should remain involved at all levels.

For SDOs, the value of data can be increased through the development and use of standards, and contribution to standards development and implementation is essential. This was illustrated with the example of the ISO IDMP standard.

For the European Commission, the EMRN data strategy is closely linked with the EHDS and its implementation. Built on the same principles for data, they both foster greater access to data as well as safe, secure and ethical use of data. While interoperability is supported in the EHDS via the metadata catalogue, strengthening locally data quality through the EMRN data strategy will also contribute to greater interoperability of data and systems.

For Ethics, characterisation of ethical options in their diversity are metadata to consider. Ethics should be embedded in training and practical ethics literacy should be included in the strategy. What is meant by 'to the highest ethical standards' and 'as open as possible, as close as necessary' in specific contexts should be defined and the possible conflict of interests also in the use of 'ethics reasons' should be documented.

Closing remarks

The event was closed by the event chair Peter Arlett (EMA, BDSG co-chair).