EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

18 January 2020
EMA/10082/2021

EU big data Stakeholder Forum

Summary report from the big data Stakeholder Forum held on 15/12/2020

Introduction

On the 15th of December the first big data Stakeholder Forum was hosted by EMA on behalf of the HMA-EMA joint Big Data Steering Group. The Forum aimed to give stakeholders an opportunity to provide feedback in regards to the HMA-EMA big data priority recommendations and perspectives on their implementation. During the forum speakers from different stakeholder groups presented what they considered priorities, opportunities and risks, and areas for collaboration and engagement.

This report summarises key points presented by stakeholders, firstly the broad themes presented across stakeholder groups, followed by points from individual stakeholder groups.

Common themes

During the forum, positive feedback was shared on the progress made since the establishment of the big data steering group in May 2020 and on the adoption of its workplan in July 2020. Given the challenges from the ongoing COVID-19 pandemic, it was appreciated that the agenda for change, initially set out was still on track. The use of health data gathered outside of clinical trials is seen as promising for novel insights into disease, as well as the development and regulation of medicines. Use of real-world data (RWD) through the DARWIN EU project is expected to improve health outcomes for all European citizens. The ability to quickly identify patients for drug trials at the EU level will improve the speed at which evidence is generated and improve responsiveness to health crises. Stakeholders noted the hurdles to implementing secondary use of RWD include data quality and representativeness, accessibility, robust study methods and establishing the evidentiary value and regulatory accessibility of evidence generated. It was proposed that a collaborative approach to addressing the challenges, including developing best practice guidance will be important. Stakeholders showed consensus in recognising the value of RWD as a complement to clinical trials.

Data quality and representativeness is a top priority for all stakeholders, and guidelines are requested to ensure quality in the context of different data sources and use cases. Stakeholders emphasise interoperability as key for efficient RWD usage, and the wider importance of adhering to FAIR principles. RWD can fill the gaps where evidence from clinical trials is lacking but can only deliver on this promise if national variations are overcome by adopting common standards on data quality and metadata.

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

All stakeholders emphasised the need for a secure and ethical governance framework. Complete data protection (GDPR) compliance is essential, and the need for governance frameworks was stated and these should enable stakeholders to know where data is kept, how it is used, and who is responsible for processing.

Stakeholders emphasised the need for intensive cross-stakeholder collaboration and enthusiasm to participate in planned workshops and consultations.

Specific comments

For patients, the ability to fulfil unmet medical needs and repurposing of drugs are priorities for using RWD. For this purpose, patients are willing to share their data for the common good. The biggest risk identified through a patient survey, was breach of confidentiality and subsequent risk of re-identification. This will require rules for pseudo/anonymisation procedures, and security standards for IT systems. Lastly, it is important that research and innovation is not only industry but also patient-driven: "Co-creation is key."

Patients and clinicians note digital literacy as a public health challenge. There can be no health digitalisation without digital and health education. Clinicians expect big data to improve outcomes, to result in a better understanding of (and definitions for) rare disease and adverse drug reactions, and to improve the patient experience. Clinicians therefore recognise the promises of big data in the clinic, with some caveats: The lack of standardisation for collecting data, and lack of harmonisation between countries regarding the analysis of the data. These problems similarly extend to no standard IT systems. Finally, the increasing digitalisation of the clinical space challenges a vital part of human care, namely the human touch, a requirement for trust. This patient-centric view was echoed by academia.

Clinicians and researchers raised methodological concerns regarding data collection such as sampling bias and incompleteness of data. Lifetime sustainability of the big data initiatives is another point they consider, specifically after the initial investment phase. Considering the varying infrastructure, knowledge, and other capabilities of research centres and hospitals across Europe, it is vital that big data initiatives do not exacerbate existing inequality. This concern about inequality was also mentioned by academics, specifically the costs for hospitals associated with the implementation and maintenance of EU wide big data initiatives. This risks compromising the representativeness of data collected, biasing better resourced clinical centres.

Researchers welcomed the availability of large amounts of quality data, and the collaborative development of "solid epidemiologic methods". In order to enhance trust in the regulatory system three pillars should be promoted: transparency, independence, and standards, in line with the priorities developed by the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP). They also mentioned the risk for confusion between data driven analysis and causal inference i.e., the difference between prediction-oriented machine learning and traditional statistical and epidemiological techniques more concerned with understanding the relationship between predictors.

Academia highlighted the ability for the initiative to optimise the use of scarce healthcare resources. Promise is seen in a continuous and federated form of learning. Academia is willing to lend their expertise in regards to this federated learning, for collaboration but also for teaching. The development of standards should build on previous initiatives based on academic-industry partnerships. Together with researchers academia mentioned the importance of conducting research not only together with, but also independently from, any market interests.

Both industry and clinicians mention the language used to discuss these topics, and how important it is that everybody is "speaking the same language". This was demonstrated by the medical device industry, who contrasted their perspective of what additionally constitutes health data/RWD with that of other stakeholders; such as data from activity trackers or grocery shopping i.e., data generated outside the healthcare system.

Industry sees the added value of DARWIN EU and requested more details on how it will work, and which stakeholders will be involved. It is proposed that all stakeholders should have access to the DARWIN EU system. The question of data quality should be addressed through multi-stakeholder collaboration. Industry offered their experience in analysing RWD, and in sharing pertinent use cases and case studies. The generics industry additionally noted that when an off-patent product is made available, its product profile will already be well characterised through RWD. Self-care industry mentioned the potential of RWD to support the switch to non-prescription use of products. All presentations from the Forum are available here.