9 November 2022
EMA/880433/2022

Summary report of the 2nd biannual Big Data Steering Group and Industry Stakeholders meeting

3 November 2022 – co-chaired by Jesper Kjaer (DKMA) and Peter Arlett (EMA), Webex

1. Welcome and opening remarks

The Co-chairs of Big Data Steering Group (BDSG) opened the meeting and welcomed participants representing the BDSG and pharmaceutical (human and veterinary), medical devices and IT technologies industries.

Objectives of this meeting were to:

1. Follow up on industry's feedback incorporation into the revised BDSG workplan 2022-2025:
    - Jesper Kjaer presented an overview of the revised BDSG workplan 2022-2025 highlighting how industry's input (received during the first meeting in May 2022) was integrated across the activities of the big data priority recommendations.
2. Inform on the delivery of the selected BDSG workplan 2022-2025 activities:
    - EMA provided updates on the DARWIN EU® establishment and the draft EU Data Quality Framework which was released for public consultation until 18 November 2022.
3. Hear industry's perspective and have open dialogue on the selected data topics:
    - Industry presented their initial feedback on the Good Practice Guide for the use of real-world metadata (released for public consultation until 16 November), provided reflections on the clinical trials raw data pilot and sought clarifications on the progress of the EHDS pilot.

EMA and industry presentations were followed by open discussions amongst the meeting participants.

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

2. Big Data priority recommendations

2.1. Data discoverability

Good Practice Guide for the use of real-world metadata for regulatory decision making

Gianmario Candore (EFPIA) presented industry's initial feedback on the Good Practice Guide for the use of real-world metadata prior to submitting it formally by the deadline of 16 November 2022:

- Industry welcomed the initiative to create a catalogue of real-world data sources and to provide guidance on its use to relevant stakeholders.
- In relation to the data collection and maintenance process, industry highlighted the importance of ensuring the information in the catalogue is comprehensive, accurate and up to date, and in this respect provided several recommendations.
- Industry encouraged to perform catalogue's functional testing involving relevant stakeholders e.g. industry.
- Other suggestions for consideration received from industry included:
  - Link and cross-reference the Good Practice Guide with the Data Quality Framework to ensure consistency in the use of concepts (i.e., data quality, data reliability) and terminology;
  - Clarify the document's scope from the geographical (EU vs non-EU) and data source type perspectives;
  - Add explanations of the metadata and proposed values for key variables and consider additional metadata such as time lag between the collection and availability of data, more granular detail on laboratory data and inclusion of genomic data sources;
  - Consider linking the catalogue to similar initiatives such as EHDEN or EHDS catalogues to harmonise definitions and avoid duplication of effort in providing information to different sources from the data holder point of view.

Comments and ideas expressed by industry were well taken and will feed into further development of the catalogue. A workshop with data holders and industry next year may be considered helpful in this context.

It was confirmed that industry representatives will be invited to test the catalogue's functionalities in 2023.

2.2. Data quality & representativeness

Data quality framework

Ana Cochino (EMA) presented on the draft version of the Data Quality Framework (DQF) by outlining the steps taken in drafting the document, introducing the fundamental data quality concepts, and informing of the public consultation process and next steps.

It was highlighted that the purpose of this framework was to be applicable for any type of human and veterinary data that might be submitted to medicines regulator. The work will continue next year to apply data quality principles outlined in the framework to specific domains (e.g. real-world, manufacturing etc.) and to ensure alignment with developments coming from the European Health Data Space.

In response to industry's question as to whether guidance should be expected on the data quality framework application and applicability of certification of data sources, it was explained that such discussions were envisaged in the next iterations of the DQF.

A suggestion was received to pilot this data quality framework with the recently onboarded DARWIN EU® data partners.

Industry stakeholders were encouraged to share the document with their data managers dealing with e.g. clinical trial or real-world data to reflect and feedback to EMA (through the public consultation process) on how the concepts outlined in the framework mapped with their data quality processes.

2.3. DARWIN EU®

DARWIN EU® update including approach for consulting/informing industry on studies

Andrej Segec (EMA) gave an update on the ten data partners that were being onboarded into the DARWIN EU® network as part of the Coordination Centre's Phase 1 of establishment, first studies to be conducted in 2022 and approach for consulting/informing industry on studies to be performed through DARWIN EU®.

Industry welcomed the concept that study complexity will be the driver for industry's consultation on DARWIN's studies and enquired whether in the early stages of the DARWIN EU establishment, case by case judgment for selected studies of other categories (i.e. non-complex studies) could be considered. Such feedback was currently being assessed by the DARWIN EU® team, also considering that by 2025 the number of studies will significantly increase and case by case consultation will not be feasible. The final proposal on industry's consultation is expected to be presented to the DARWIN EU® Advisory Board in December 2022.

Industry stakeholders were also informed of the upcoming press release on DARWIN EU® data partners in November 2022.

2.4. EU network capability to analyse

Clinical Trials raw data pilot

Almath Spooner (EFPIA and industry representative of the Raw Data Industry Focus Group) expressed appreciation of the ongoing dialogue with the raw data project team via the designated industry focus group and in this respect provided the following reflections:

- Industry had experience from submissions of raw data to FDA and PMDA and practical aspects as well as learnings from such submissions were being shared through the focus group. However, context for the submission of raw data is different in the EU due to different disclosure and data privacy policies.
- It is important that transparency and predictability for sponsors is provided in order to understand how the data they submit will be analysed and integrated into medicines assessment procedures as well as have opportunities to offer comments.
- Further elaboration is required on how pilot's benefits will be evaluated and what metrics will be used during the pilot to determine value of a new approach on the use of raw data for decision-making.
- Resourcing models: how will the pilot evaluate current data analysis capacity in the regulatory network and inform future capacity and capability building.

- Industry expressed the need to strengthen the dialogue between the regulatory network and industry relevant experts to clarify data protection and transparency concerns.

Meeting participants were re-assured that EMA and regulatory network are committed to work through the details of the data protection and transparency aspects during the pilot and a dedicated discussion with the pilot industry focus group on this subject will take place.

Governance framework

EHDS2 pilot

Inka Heikkinen (EuropaBio) opened the discussion by posing a few questions related to the progress of the EHDS2 pilot and how it will interface with DARWIN EU®.

It was explained that the regulatory network's purpose in the EHDS2 pilot was to contribute to the overall building of the EHDS as well as understand and assess the EHDS impact on medicines regulation in general. In this context, an EMA use case through DARWIN EU® has been included in the EHDS2 pilot and the results of this use case were expected in 2024.

The pilot for EHDS is led by the French Health Data Hub and possibility of inviting their representative to a future meeting with industry stakeholders will be considered.

3. Wrap-up and conclusions

The BDSG co-chairs thanked the meeting attendees for their participation and concluded on the importance of continuing active dialogue and collaborative work with stakeholders in our journey towards the vision for the data driven medicines regulation.