Technical workshop on real-world metadata for regulatory purposes – Opening remarks

Technical workshop on real-world metadata for regulatory purposes
Virtual meeting, April 12, 2021

Presented by Nikolai Brun, DKMA
Co-chair, HMA-EMA joint Big Data Steering Group

Workshop objective

The objective of this workshop is to review and gather stakeholders' feedback on:

o Preliminary list of metadata required for characterising real-world data sources and definitions to fulfil regulatory use cases;

o Proposed options for metadata from real-world data sources collection and maintenance processes;

o Proof-of-concept catalogue of data sources and metadata.

The Big Data Task Force report from 2020 Sets the stage

Recommendation #3:

Enable data discoverability. Identify key **metadata** for regulatory decision-making on the choice of data source, strengthen the current ENCePP resources database to signpost to the most appropriate data, and promote the use of the FAIR principles (Findable, Accessible, Interoperable and Reusable).

Big Data Steering Group: 2020 Report

https://www.ema.europa.eu/en/documents/report/big-data-steering-group-bdsg-2020-report_en.pdf

Established project, funding and operating model + initiated planning for pilot of European Health Data Space

Defined scope for a study on a data quality framework for medicines regulation

Initiated a study to establish 'meta-data for real world data (RWD)'

Delivered a training signpost, a real-world evidence (RWE) curriculum and biostatistics curriculum

Progressed a PRAC RWE analysis pilot + initiated a CHMP review of RWE in MA submissions

Initiated a CHMP pre-pilot of CT raw data analysis + selected software for RWD analytics

Initiated review of methods domain of EMA working parties

Delivered data protection workshops + initiated a data protection Q&A

Initiated analysis for an international collaboration roadmap on RWE

Delivered a multi-stakeholder forum in December 2020

Established a veterinary data workstream

Big data workplan: 2021

1. DARWIN	Coordinating centre establishment initiated | DARWIN Advisory Board established | Pilot of the European Health Data Space initiated
2. Data Quality	Initiate Data quality and representativeness study
3. Discoverability	Initiate enhancement of EU catalogue of real world data resources | Best practice guide on Metadata for regulatory purposes | Meta-data workshop 12 April 2021
4. Skills	Training curricula published | deliver at least one module per curricula (stats, epidemiology, data science)
5. Processes and transparency	Publish learnings from review of RWE submissions + learnings from committee RWE analytics pilot
6. Analytics capability	Patient level data from clinical trials: pre-pilot becomes pilot | AI Workshop 19/20 April 2021
7. Expert advice	RWE and advanced analytics expert advice available
8. Ethics and Security	Publish EMA Q&A on secondary use of health care data and data protection
9. International	Publish international collaboration roadmap on RWE | Standardisation workshop May 2021
10. Stakeholder forum	Stakeholder forum December 2021
11 Veterinary	Agreement on applicability of the BDTF recommendations to the Veterinary domain | Vet workshop 1 and 2 June

BDTF report on Metadata: (section 5.4 on Data Discoverability)

For Big Data approaches it is essential that automated data processing algorithms efficiently identify appropriate data sets, based on the information provided by the **metadata** of a respective data set.

The **metadata** have to have a format that is 'readable' by the used data processing algor

FAIR principles suggest that **metadata** should be generous and extensive, and should include information about the context, quality, and condition, or characteristics of the data and should not pre-suppose a purpose or user for the data.

US FDA and Japanese PMDA: Both agencies mandate CDISC standards for datasets and associated **metadata** for marketing authorisation applications

Thank you