30 August 2023
EMA/EMA/392935/2023
Data Analytics and Methods

Report from "Multi-stakeholder workshop on Real World Data (RWD) quality and Real World Evidence (RWE) use"

26-27 June 2023, hybrid meeting, EMA, Amsterdam

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

Day 1 - Monday, 26 June 2023

Welcome and opening speech

Peter Arlett, European Medicines Agency (EMA), opened the meeting by welcoming everyone to the workshop. He stressed the importance of working with the joint action Towards the European Health Data Space (TEHDAS) and leveraging different groups of stakeholders to enable a transformative approach to generating clinical (and in particular real-world) evidence with a standard of excellence. Profound changes ahead regarding science and technology, e.g. Artificial Intelligence (AI) will bring value to public health, but we need to ensure we address its challenges and establish its rigor.

Sara Almeida, European Commission (EC), then introduced the EU Pharmaceutical Reform which proposes a streamlined and agile regulatory framework catering for innovation. She highlighted the accomplishments of the DARWIN network including the 27 studies and the onboarding of many new partners and reiterated the Commission's commitment to enabling the use of RWD and having a regulatory framework that support it.

Jesper Kjær, Danish Medicines Agency (DKMA), and the Heads of Medicines Agency (HMA) Big Data Steering Group (BDSG) co-chair then pointed out the importance of fostering better use of RWD also at a national level, and to share use cases from different member states. Processes, technologies and data access agreements must be optimised at a local level to enable a better functioning European evidence network. The Data Quality Framework (DQF) will be important to ensure we have a harmonised understanding of data quality.

Session 1: Development of the Data Quality Framework – status update and next steps

TEHDAS and EMA have both drafted Data Quality Frameworks in close collaboration to ensure their complementarity.

The Data Quality Framework is the flagship work from the TEHDAS group to be released in July 2023 and came out of an effort for compromise between the scientific/theoretical considerations for data quality and its implementation as "utility" within countries. Its aim is to provide solutions for the trustworthy secondary use of health data to promote the digital transformation of the European health system by outlining key elements that data holders should consider enabling data "re-use". Those include data quality management and quality assurance, approach to semantic interoperability, dataset publication and cataloguing, minimisation and purpose limitation and governance. The framework also contains guidance for the implementation in member states.

Drafting of the EMA Data Quality Framework started in 2022 under the recommendation from the BDSG and was sponsored by the Methodological Working Party (MWP) as well. After receiving more than 500 comments from various stakeholders during the public consultation on the scope, the need for further deep-dives, and some specific technical points, the final DQF will be published imminently (summer 2023). A deep-dive chapter on RWD is now being drafted, which this workshop will help inform and which will ultimately be subject to public consultation.

Q&A

The need for tools to support the adoption of the EMA DQF RWD deep-dive was highlighted. The connection of the two frameworks to existing standards like DCAT (Data Catalogue Vocabulary) was discussed. Explanations were given regarding the aspiration to move data sources towards semantic interoperability and adoption of Common Data Models in an agnostic way, i.e. there are other options

than just OMOP (Observational Medical Outcomes Partnership) specifically.
It was highlighted that for TEHDAS, part of assessing the fitness-for-purpose of a specific data source is enabled through users providing feedback on their use of a data source for a specific purpose. However, we are still far away from a "learning healthcare system" in the EU.

Session 2: Use cases: approach for assessing and improving Data Quality

FIMEA (Finnish Medicines Agency) serves as a national competent authority and also functions as a registry holder which enables access to it's data sets. FIMEA also generates and utilize Real-World Evidence (RWE). Finland has a national metadata catalogue, and FIMEA contributes by publishing its data resource descriptions there. The metadata catalogue provides data resource descriptors in an overall level down to variable and code levels. This catalogue also includes information on data quality. Both, a metadata catalogue and clear data management processes, are important for data quality at institutional level.

A use case/pilot, focused on federated analysis, was presented involving three university hospitals acting as data partners. In the pilot FIMEA defines research questions, pools aggregated statistical data and informs the data partners about data quality requirements. Some identified examples of data quality indicators were identified from EMA's report on CAR-T cell therapy Registries Workshop.

There are over 20 registers managed and stored by Danish Health Data Authority, covering the entire national population for over 40 years. The National Patient Register can help answer many research questions, especially as it is possible to link all health data through a personal identification number.

Data quality in Danish Health Registers is fundamentally based on legislation – all hospitals are obliged to report patient treatments to the National Patient Register (assurance on coverage/volume), clear terminology (high quality due to standardisation and digitation, common standards that are being maintained by the Danish Health Data Authority), classifications (Danish classification based on WHO but also national extensions), guidelines and technical validation (semantic but also with description of technical matters; 467 rules for validation which are used as guidelines for technical partners in hospitals), governance and co-operation within the health sector (boards and committees with meetings for better use of health data), culture and tradition. For the future, AI is expected to be leveraged to analyse the quality of data in a more automated way.

The Health Data Hub is a public interest group aiming to facilitate access of secondary data, and its data catalogue includes one of the world's largest healthcare claims databases (SNDS) which is a large claims database covering virtually the entire French population. In France, there is an increasingly important role of RWD in access to reimbursement and pricing – the complementarity of RWD to clinical trials is now recognized at all stages of evaluation. Focus areas within the Health Data Hub are:

1. Organise and promote the collection of RWD: create simple, clear and secure governance rules. Make data available quickly. Provide reference methodologies for the collection and use of harmonized data.
2. Enhance quality and interoperability: e.g. SNDS turning to OMOP, increase quality through extensive standardisation of data collection, processing, terminology and design principles.
3. Develop innovative methodologies: e.g. creation of synthetic control arms , AI for drug discovery.

Dr Kelly H. Zou, presented on industry's perspectives on behalf of seven industry groups. Identifying fit-for-purpose data sources entails identifying the right type of RWD sources (e.g., electronic health records, claims, etc.), the appropriate periods that such RWD cover, purpose and criteria for data collection, feasibility assessments with queries on patient counts, missing data, data generation

details, and past use-case examples from data holders. Existing tools and approaches may be considered and leveraged whenever possible, such as those through the Observational Health Data Sciences and Informatics (OHDSI), tools like ACHILLES and the Data Quality Dashboard, REQueST, SPID, TransCelerate's RWD Audit Readiness Considerations document, Authentic Transparent Relevant Accurate Track-Record (ATRAcTR), and more. Building efficiency and best practices will rely on sponsors' and data holder's willingness to participate, their comprehension of the data quality framework, and the availability of data sources and metadata catalogue.

Q&A

It was pointed out that there are opt-out possibilities, but it is very rare that they are used – patients usually do want their data to be used. The need to increase the ability of registries for augmenting or adapting their data quality in an agile/timely manner was discussed; the Finnish pilot project looks forward to creating more of a feedback loop in discussions with the registry holders. The need to consider how EU-level frameworks and national tools and pilots will work together was raised.

Session 3: Systems and processes underpinning Real-World Data

The section of the EMA DQF RWD deep-dive regarding "foundational determinants", i.e. what contributes to the quality of the dataset in terms of processes and systems, was presented. RWD is collected, ingested, processed, manipulated and enriched in a way that it is subject to quality issues at each of those steps. Characterising those various steps according to different maturity models – from simple documentation to formalisation/standardisation, to automated integration of metadata elements – increases the trustworthiness of the data. A checklist outlining those steps and related expected documentation is being drafted in the DQF RWD deep-dive, and maps back to most of the elements from the EMA-HMA catalogue of data sources, which is under development.

Open discussion

Industry stakeholders highlighted that they welcomed the possibility of a checklist which would clarify the level of detail expected, and suggested defining the most critical information to minimize the burden on data holders. There was broader acknowledgment from stakeholders that there is a need to consider the return on effort, and to start with a simple approach that could be iterated on.

It was clarified that the intent of the framework and its deep-dive are not to provide criteria to provide qualification of a data source, but rather to outline what evidence should be available to perform a fitness-for-purpose assessment. It was also clarified that the framework does not intend to impact primary data collection directly.

Session 4: Data Quality metrics for Real-World Data

DARWIN EU® is a federated network using the OMOP CDM, which uses the Kahn data quality framework as a basis for its quality assessment, in line with the EMA DQF. A 3x2 matrix with 2 checking procedures (verification and validation) vs 3 quality dimensions (conformance, completeness, plausibility) is used as a framework for identifying metrics to perform large-scale, systematic assessment of data quality. This has resulted in more than 4000 checks, most of which are verification, which are applied via the Data Quality Dashboard. Thresholds have been established for how many records can violate a particular check, and during onboarding meetings with the data partners, data quality issues that were flagged (i.e., above threshold) are discussed which is helpful for improvement and understanding the explanation for these issues.

For the EMA DQF RWD deep-dive, a slightly different framework to DARWIN EU® for identifying data quality metrics is proposed, with checks not just on completeness, plausibility and conformance, but also checks in comparison to other datasets, checks on dataset descriptors and objective checks covering more than just completeness. There are maturity levels in the way that these checks are being performed and reported, from self-reporting (qualitative) to quantitative application of metrics and ultimately a feedback loop. Considerations regarding the application of metrics were highlighted including: metrics can be applied at different points in the lifecycle, the result can change for a subset of data of interest, the responsibility for generating metrics must rely with different actors throughout chain of evidence, a large number of tests can be difficult to interpret and the EMA-HMA data sources catalogue can serve to capture data quality metrics.

Open discussion

Clarifications were brought regarding the DARWIN EU® dashboard: thresholds are not set for a particular research question, they just serve for general onboarding but can be tweaked/selected based on the context. There is not a yes/no assessment based on the results of the metrics test, instead there is a manual review of the tests for which there were large numbers of failed records. It is difficult to automate the assessment of fitness-for-purpose of data sources as that should still lie with the researcher who should consider that RWD will always have a certain degree of error. Comparing data quality across data sources is valuable – while we cannot define a "right answer", we can expect that there should not be too big a divergence between sources.

Session 5: Data Quality in the context of a regulatory/research question

The purpose of conducting a fitness-for-purpose data quality assessment is to gain the trust of the end-user regarding study results. The ENCePP checklist for study protocols helps understand how researchers initially choose a study design, then should report on data elements and their validity. Examples were given where showing how one might fulfil different study objectives and the data requirements associated with it. For example, for a confounder assessment, if some of the confounders are missing in the data source, one would need to think about how to address that (e.g., through imputation) as well as assess the impact of that data missing.

Decision-makers (end users of study results) may not be familiar with RWD sources included in a study, so one needs to consider how to ensure a level of trust. Most essential to that are reliability and relevance. For reliability, it is important to reassess foundational and intrinsic checks for specific research question to understand the primary purpose of data collection, the representativity of data for inferential purposes, the availability of peer-review publications of the data source, and metrics to increase confidence in the data source. For relevance, it is important to understand if the right data is available at the right level, i.e., do we have the data within the computational phenotypes of interest (e.g., disease codes, medicine names, etc). A framework inspired from the SPIFD (Structured Process to Identify Fit-For-Purpose Data) is being developed in the EMA DQF RWD deep-dive that could be used to perform that assessment.

Open discussion

It was highlighted that study design should not only be informed by research question, but it may also be guided by the constraints of the data and may be adjusted based on results from a feasibility study.

Stakeholders discussed that scoring against a "fitness-for-purpose assessment matrix" may not be the end goal but can also be a roadmap for improving data collection itself.

Concluding remarks

Jesper Kjær and Peter Arlett, on behalf of the BDSG, thanked all the presenters and moderators of the workshop, and emphasised the need to continue efforts regarding RWE and Data Quality with a multi-stakeholder approach.

Post meeting note:

EMA would like to acknowledge IQVIA's support in drafting materials and organising of this workshop first day, as well as the continuous support throughout the drafting of the data quality framework. A special thanks to Andrea Splendiani, Annabelle Monnot, Hanne Van Ballegooijen and Jamie Skipper.

Day 2 - Tuesday, 27 June 2023

Session 1: RWE in regulatory assessment and decision-making processes.

Stefanie Prilla (EMA) presented the results of the EMA review of the experience gained with conducting RWD studies to support EU regulatory decisions. The review focused on studies that have been conducted in the context of ongoing pilots with several of EMA's scientific committees and the Scientific Advice Working Party (SAWP). Additional studies performed in response to the COVID-19 public health emergency and the Pharmacovigilance impact strategy were not included in the main analyses.

The report follows the priority recommendations of the Big Data Task Force as implemented through the BDSG and the second multiannual work plan (2022-2025). It covers the period from September 2021 up to 7 February 2023. During this period, a total of 61 RWD research topics were identified, and 30 studies were initiated. There was a broad range of study types including safety studies, drug utilisation studies, disease epidemiology and clinical management studies, and studies to inform the design and feasibility of clinical trials. The studies were able to address a range of research questions and help support decision-making in a variety of regulatory contexts and procedures. The report points out the need for wider access to additional, more diverse, and complementary data sources, including hospital, claims and registry data. Similarly, data sources from additional European countries are desirable to broaden geographical representativeness. The report also highlights the need to accelerate the generation of RWE, which is planned to be achieved with DARWIN EU® by the development of a catalogue of standard data analytics that can be readily executed, of phenotype libraries, and of precomputed dashboards. In addition, the early identification and better anticipation of possible research questions might allow for the conduct of more complex and thus time-consuming analyses despite short procedural timelines. Close collaboration with decision-makers and other stakeholders has shown to be crucial for the implementation of a fit-for-purpose RWE framework. The learnings and recommendations emerged from the review will feed into the work of the BDSG and further inform the establishment of EMA RWE framework. For more information please refer to: RWD studies report | Review: real-world data studies (info sheet).

Álmath Spooner, European Federation of Pharmaceutical Industries and Associations (EFPIA), presented an overview of use of RWE by the pharmaceutical industry, underlining that industry recognises the role of both randomised controlled trials (RCTs) and RWE as providing complementary evidence to support medical product development and regulatory submissions. Several examples were provided of how fit-for-purpose RWE has been evolving beyond traditional use cases and how its value is increasingly been recognised. For instance, during the COVID-19 pandemic, RWE provided a timely

mechanism to address uncertainties from RCTs about the effectiveness of COVID-19 vaccines in specific subpopulations (e.g., people aged 60+ years). In the case of new product approval, the use of an external control arm can help establish the magnitude of the benefit and contextualize the results from single-arm clinical trials. In this regard, early dialogue with EMA through the PRIority MEdicines (PRIME) scheme and scientific advice has proved to be valuable. Other examples included the use of RWD in combination with clinical trial to support accelerated regulatory approvals in situations of rare diseases, poor prognosis or limited therapeutic options available. In oncology, some regulatory authorities have accepted the use of RWD as supportive information to get regulatory approval in ultra rare indication and lack of standard of care (absence of consensus on the most appropriate chemotherapy).

However, global regulatory authorities are at varying stages of evaluation, developing and implementing policies for RWE and there are opportunities for international collaboration, with a view to getting convergence of guidance and best practices. Importance of consistency through good scientific practices and sound research methodology, alignment on fitness for purpose of data, better understanding of dimensions of uncertainty depending on context were highlighted. The conclusion was that medical product development is a global effort which will benefit from an increased international convergence on acceptability of RWE.

Carla Torre (University of Lisbon) presented a regulatory perspective on the use of RWE in medicines development and regulatory decision making, its challenges and opportunities. Carla underlined that the use of RWE to support decision making is not new. The landscape of medicine development has evolved with an increasing number of medicines for rare diseases and advanced therapies which do not align with the traditional drug development pathway as RCTs may be unfeasible or unethical. For instance, the COVID-19 pandemic offered a window to leverage RWD to inform clinical and regulatory decisions. However, while several well-conducted studies using RWD were crucial for decision making (e.g., COVID-19 vaccine effectiveness in subgroups, COVID-19 treatment during pregnancy). some studies used inappropriate methods which generated misleading conclusions and led to several publication retractions/withdrawals. Trends in the use of RWD were also reported for recent medicine approvals. While recent reviews showed wide-spread use of RWD in regulatory submissions, it is difficult to isolate its exact impact on decision making as regulators always consider the totality of evidence available.

The use and reliance on RWD are gradually increasing as its potential to support regulatory decision on benefits and risks of medicines throughout their lifecycle is being unlocked. The presentation highlighted a number of challenges, solutions and opportunities:

- Data, namely, the heterogeneity of data source types, data quality, and data governance, sharing and access. Reliable RWE is built on use of fit-for-purpose RWD, i.e., data that speak to the question at hand and are of high quality.
- Methods, which should consider analytical integrity (i.e., appropriateness of study design and data analysis) and compliance to the best methodological standards e.g., European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) guide on methodological standards for pharmacoepidemiology, EMA Good pharmacovigilance practices (GVP), etc.
- Trust, which should be increased by improving transparency on study designs (pre-registration of protocols), as well as reporting with responsible communication of results.
- Policy and governance environment, for which it is important to strengthen international collaboration to work towards comprehensive guidance on the use of RWD, common definitions,

best practices, capacity building and continuous engagement with all stakeholders (including patients and healthcare professionals).

Final remarks emphasized that RCTs and RWE are complementary for drug development and decision making as opposed to being competing alternatives. The role of RWE in supporting medicine evaluation has evolved over time, however, enabling its use and establishing its value will require further work. Early dialogue and frequent interactions with regulators are key for the generation of fit-for-purpose RWE.

Stefanie Prilla (EMA) presented the results of a survey conducted amongst participants of the workshop to get a snapshot of the different views on opportunities and challenges to fully integrate RWD/RWE in regulatory decision-making. 38 participants from four groups responded: 1) regulators and health technology assessment (HTA) bodies, 2) pharmaceutical industry, 3) healthcare professionals (HCPs), learned societies, patients and non-governmental organisations (NGOs), and 4) academia, researchers and data providers. The survey showed different perceptions of responder groups with respect to the current relevance of RWE for decision-making. While most of the responders from pharmaceutical industry and academia/data providers considered the use of RWD essential because it can fill important gaps on top of Clinical Trial (CT) data, regulators/HTA and HCPs/patients see RWD mostly important for selected regulatory purposes. Responders across the four groups regarded RWD as most suitable to inform on the safety of medicinal products and other established areas of use such as disease epidemiology and drug utilisation. Less established use case categories, such as effectiveness of medicinal products, understanding feasibility of clinical trials and context to single arm trials also ranked high for responders from the pharmaceutical industry. When asked about the main challenges to fully integrate RWD/RWE in regulatory decision making reported were, most responders chose: insufficient information in existing data sources, difficult access to data, and limited acceptances of RWD/RWE by decision makers. In relation to the question what needs to be done to fully enable the use of RWE in regulatory decision making, development of guidance on best practice in RWE, increasing transparency on relevant data source characteristics, and building trust in the use of RWE for regulatory decisions was chosen by most responders. Most responders considered the role of regulators in generating RWE useful, especially in view of using/integrating national data sources through a network such as DARWIN EU®, and important as it allows for independent RWE generation. When asked about how the use of RWD would be expected to evolve in the coming years, most responders considered that the use will grow including in areas less established.

Panel discussion

Jorge Batista, Pharmaceutical Group of the European Union (PGEU), started the panel discussion with some reflections on the involvement of pharmacists in RWE generation. Pharmacists are already engaged in RWD collection and use. For instance, pharmacists contribute to the reporting of adverse events, or data collected through electronic health records. Such data are used for pharmacovigilance in intensive monitoring studies, assessing adherence, evaluating shortages, or for post-authorisation studies on safety and effectiveness. Pharmacists also collect data on drug use and quality of life in collaboration with other organisations (e.g., HTA bodies). Therefore, envisioning the collection of RWD at European level will require open communication with all the stakeholders, training, and capacity building to raise awareness on reasons for the data being stored and used, including data privacy concerns.

The panel discussion focused on priority actions that need to be undertaken to facilitate the use of RWE in regulatory decision making. Key messages were:

- Medicine developers would welcome being part of the dialogue to build regulatory guidance on conduct of non-interventional (observational) studies.
- Early dialogue and engagement of medicine developers with regulators through SAWP has shown to be helpful for a successful approval process.
- RWD is useful to inform clinical trials, however acceptance needs to be enhanced by improving methodology, terminology, validation of data sources.
- Further RWE reports should include studies conducted to support marketing authorisation applications (MAA) which could provide a more comprehensive picture.

Finally, Michel Berntgen (EMA) asked the panelists to choose a single action that should be undertaken collectively to facilitate the use of RWD in decision making. To which, panelists replied: engagement (pharmacist community), predictability – i.e., fitness for purpose data sources, use cases - and alignment (industry), building trust (regulators). Main answers from the audience were aligned with panelists choices: trust, transparency, collaboration, guidance, quality, harmonisation and dialogue.

Session 2: DARWIN EU® where we are in the Phase 2 of its implementation.

Andrej Segec (EMA) provided an overview of current state of the implementation of DARWIN EU®. During the first year of its implementation, 10 data partners were onboarded, and four studies were initiated. The second phase of implementation is in progress with 16 studies (including different use cases across the medicinal product lifecycle), the selection of further 10 data partners, and the establishment of analytical pipelines and codes. Study protocols and reports of ongoing and finished studies can be found in EU PAS register. The DARWIN EU® Coordination Centre launched its website (www.darwin-eu.org), which contains detailed information on a catalogue of standard data analyses for off-the-shelf and complex studies that can be conducted. This catalogue is currently under consultation involving the MWP, DARWIN EU® Advisory Board, and industry.

Juan Jose Abellan (EMA) presented the outcomes of the EMA virtual workshop held in October 2022 with the participation of 30+ representatives from HTA and payer organisations across the EU. The objectives of the workshop were to raise awareness on the possibilities of RWE generation via DARWIN EU®, to better understand HTA/payers possible research questions, and to identify use cases from HTA/payers perspective that could lead to studies conducted by DARWIN EU®. Key messages from the workshop with HTA/payers were:

- the need to address concerns with regard to the quality of RWD to further enable the use for decision making;
- the need to promote further awareness on how DARWIN EU® operates, the type of data available via DARWIN EU® (which can affect relevant questions to be addressed);
- the need for studies on effectiveness, natural history of disease, and for use of appropriate methodologies (e.g., use of digital data, AI and Machine Learning).

Regarding uses cases for possible studies, several suggestions were discussed. The following two were selected post-workshop in collaboration with HTA/payers representatives:

- Study on the natural history of multiple myeloma to characterise patients with multiple myeloma, treatments received (monotherapies, combinations) including sequences, and survival.
- Study to characterise patients with non-small cell lung cancer treated with immunotherapies (pembrolizumab, nivolumab, atezolizumab, etc.) as first line.

Mario Jendrossek (Health Data Hub) presented an overview of the European Health Data Space (EHDS), identified by the European Commission as a top priority of the European health policy, and its pilot. The project will cover important pieces of the overall EHDS use journey from data discovery, data permit requests, data preparation, use of data and report (i.e., publication of the study and valorisation).

Five concrete research use cases of growing complexity were selected to demonstrate the feasibility and added value of European research projects. Of these, one is led by EMA via DARWIN EU® on identifying risks of coagulation disorders in patients with COVID-19. Expected deliverables include high-level research protocol, analysis scripts (for OMOP and for native data), input to the project work packages, and publication in the EU PAS register.

Erika Duffell, European Centre for Disease Prevention and Control (ECDC), introduced the global initiative to scale up the attention to hepatitis as a public threat and how ECDC is working on supporting countries in addressing the global goal of viral hepatitis elimination by 2030. There is still a substantial burden of hepatitis B and C in Europe. To achieve the goals of reduction of the number of new infections and deaths by 90% and 65% respectively, there are some challenges that need to be tackled including gaps in availability of national-level data and data for key population subgroups (particularly regarding treatment), heterogeneity in data collection and quality across monitoring systems and data networks which affects comparability between countries, and data recency.

The collaboration between ECDC and DARWIN EU® seeks to explore how timely data on hepatitis treatment from several countries can be generated to inform/support the monitoring of progress towards the hepatitis elimination targets in Europe, particularly with a focus on key population groups.

Panel discussion

Katia Verhamme, DARWIN EU® Coordination Centre (CC) joined the panel discussion and introduced the DARWIN EU® CC and the progress made during the first year on establishment as well as on the conduct of the first four studies (3 off-the-shelf and 1 complex studies). The interest of the DARWIN EU® CC to understand the needs from various EMA stakeholders, including ECDC and HTA/payers, and to collaborate closely with them was highlighted.

Key messages of the discussion were:

- DARWIN EU® CC obtains estimates from several regions and discusses the results, taking into account particularities of each health system. DARWIN EU® CC works with data partners which are onboarded based on defined criteria (e.g., quality of data, capability to map data to OMOP CDM…). Representativeness of the data is also considered, and this aspect will improve as more data partners are being onboarded. Inclusion of all European countries and all healthcare settings is a challenge; however, DARWIN EU® CC plans to have onboarded 40 data partners by 2025.
- DARWIN EU® CC publishes study protocols and reports in EU PAS Register once these are finalised.
- DARWIN EU® CC consults industry for complex studies at the stage of the study protocol development. The relevant marketing authorisation holders (MAHs) will receive notification in advance that they will be consulted.
- Complex and very complex studies require more work in terms of phenotyping and validation. DARWIN EU® CC also considers during the feasibility assessment step data sources in which previous validation has been done.

- Clear and simpler guidelines for sharing data are needed given the current complex rules and governance restrictions. When selecting data partners, DARWIN EU® CC considers the process for institutional review board approvals, including the possibility of getting umbrella approvals which allows for similar studies to be automatically approved and help to speed up the process.
- The heterogeneity across data sources is intrinsic to the EU given the different healthcare systems, different reimbursement processes, etc. DARWIN EU® CC considers important to understand the sources of heterogeneity to contextualise the use and help to interpret the results more accurately.
- There is a need for convergence in terminology and methodological guidance for RWE as it already exists for CTs. In addition to international efforts via International Coalition of Medicines Regulatory Authorities (ICMRA) and the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH), EMA is planning a workshop in autumn 2023 with a focus on methodology.

Session 3: Opportunities and challenges related to the use of registries

Patricia McGettigan, the Pharmacovigilance Risk Assessment Committee (PRAC), presented an overview of the use of patient registry data in regulatory decision making. The Patient Registry initiative was established in 2015, based on the need for information across the lifecycle of medicinal products beyond data coming from clinical trials (e.g., information to better understand disease characteristics, disease progression, current clinical care, safety and effectiveness of medicines). Registries that can provide evidence "usable" for regulatory purposes should bring adequate population coverage and duration of follow-up, be with standardised and timely data collection, and of high-level quality. Exploration of how registry data can be leveraged throughout the product lifecycle should start early on during the medicinal products' development, e.g. to understand the diseases natural history, to complement data collected through clinical trials, as well as to establish the best way to collect long term effectiveness and safety data post-approval. The latter will support the monitoring of the products to ensure they continue to be effective and safe, as well as to support further development such as extensions of indications. However, the use of registry data brings some challenges related e.g. to the data quality, accessibility/availability, and lack of linkage to routinely collected healthcare data, which need to be addressed to inform regulatory decision making. To this end, registries can adapt (as highlighted in the example of cystic fibrosis based on the Cystic Fibrosis Foundation Patient Registry - CFFPR). At the time of initial marketing authorisation application for Kaftrio (ivacaftor / tezacaftor / elexacaftor) in patients with specific gene mutations, registry data were considered inadequate to inform on efficacy in different genotypes. However, by the time of a subsequent extension of indication application, registries had addressed the gap by including genotype data in their collection as well as more informative data on clinical endpoints to inform regulatory decision making.

There is currently no systematic recording of the nature of the data successfully supporting regulatory applications in assessment reports, which is important in order to understand regulators' evidentiary needs. Highlighting cases in a structured approach to describe nature, contribution, strengths, limitations of supporting data would help to learn more about the use of registry data in regulatory decision-making. This could be achieved via a more structured approach in the assessment reports: a standardised summary describing the nature, contribution, strengths and limitations of supporting data would be very useful to evaluate data characteristics and quality.

Pamela Dobay and Meritxell Sabidó (EFPIA) presented a use case with no possible direct access to data and where the REQUeST tool designed by European Network for Health Technology Assessment (EUnetHTA) and its revised version published on the EMA guideline on registry-based studies was applied to assess the EMA DQF principles. The data quality dimensions and metrics used for evaluation

were reliability, extensiveness, coherence, and timeliness. Important learnings from this exercise were highlighted and three suggestions were made to EMA to improve DQ assessments and documentation to implement the DQF:

- Continue leading dialogues to set clear expectations for DQ, assessments, processes, and documentation as well as to maximize the utility of registry data for regulatory purposes.
- Consider guidance on minimum information that should be readily available to improve efficiency and transparency of DQ assessments.
- Lead the co-development and piloting of tools aligned with the DQF with all stakeholders.

Lars Wallentin, European Society of Cardiology (ESC), presented the design of a registry based randomised clinical trial (R-RCT): the Quality Registry SWEDEHEART is a web-based registry study with all data registered online directly by the caregivers. It contains standardized structured data on patients, treatments and outcomes. It is supported by an organisation of health care professionals, researchers, patient representatives. An example was shown on how data from this registry was used to follow treatments in ST-segment elevation myocardial infarction (STEMI) over time and the influence on mortality. The results demonstrated that the prolonged survival in patients with ST-elevation myocardial infarction during the last 20 years is related to the implementation of evidence-based treatments. This initiative led to the establishment of an international collaboration (EuroHeart) that provides an IT infrastructure for continuous online registration of high quality and harmonised patient-level data with real-time information supporting continuous improvement of care and outcomes in patients with common cardiovascular diseases. The project will also provide an international infrastructure for the safety surveillance of new drugs and devices and registry-based RCTs in a patient population across multiple geographies.

Kelly Plueschke (EMA) presented the results of a short survey launched in May 2023 regarding the EMA guideline on registry-based studies which was adopted and published in October 2021. The objectives of the survey were to assess awareness of the stakeholders, to identify specific topics requiring clarification, and to understand training needs.

The key results were as follow:

- 111 participants responded and most of them (76%) indicated to be aware of the guideline.
- The top five sections of the guideline considered useful were: planning a registry-based study, checklist for evaluating the suitability of registry-based studies, study protocol, data quality and data collection.
- The following sections were considered as requiring more clarity: data quality, legal obligation and regulatory requirements, and safety reporting.
- The patient experience data was highlighted by patients responding to the survey: the patient experience needs to be captured and integrated in decision making.
- The majority of respondents from regulatory, industry and researcher institutions have applied the guideline principles and recommendations.
- Lack of harmonisation, lack of interoperability, policies on data access and data sharing were seen as the main barriers for the implementation of the guideline.
- To make the guideline more visible and prominent, the following suggestions for communication and engagement were received: workshop, webinars, educational videos, Q&A documents and checklists covering fitness for purpose topics.

- Priority topics for future workshop were listed: fitness for purpose, establishing the value of registries.

The main take home message is the need for further communication and engagement, with attention to registry holders. In Q1 2024, EMA will hold a multi-stakeholder workshop on registries.

Panel discussion

Mencía de Lemus, (patient representative within Committee for Advanced Therapies (CAT), joined the panel discussion underlying the relevance for patients of the data collected post-authorisation to address remaining uncertainties. RWE can have an impact on patients' life. The challenge is to assess if the data is fit for purpose and relevant to patients.

The key message from the panel discussion was that there is a need for registries to be flexible in collecting new data elements as new therapies are made available. For instance, approvals of disease-modifying treatments for rare diseases might modify phenotypes and disease progression; flexibility is needed here to ensure those changes are captured.

Session 4: Using RWE to address public health emergencies – learnings and a view to the future.

Marco Cavaleri (EMA) introduced the session providing an overview of lessons learned from the COVID-19 pandemic on the collection of RWD and generation of RWE to inform public health crisis, in which it was essential to fill evidence gaps such as on vaccine effectiveness according to new emerging variants, number of boosters needed, and special populations (i.e., not included in the clinical trials). Key lessons learned during the COVID-19 pandemic were the need for near real-time data, collaboration between public health and regulatory authorities, and preparedness for other future health crisis. The pandemic also participated to the establishment of the Vaccine Monitoring Platform. For more information please refer to: Vaccine Monitoring Platform | European Medicines Agency (europa.eu)

Mathijs Goossens (EMA) provided an overview of the EMA-funded studies on COVID-19 vaccines to support public health and regulatory decision-making and of the establishment of the Vaccine Monitoring Platform (VMP) as a joint platform between EMA and ECDC. At the beginning of the COVID-19 pandemic, EMA worked on generating pre-pandemic background incidence rates of adverse events of special interests in preparation for the approval of COVID-19 vaccines and immunisation campaigns. Additionally, once the vaccines were approved, studies were coordinated by EMA for the post-authorisation monitoring of the safety and effectiveness of COVID-19 vaccines.

Lessons learned included:

- The continuous need for coverage, safety and effectiveness data on COVID-19 vaccines given the changing epidemiology and use of vaccines, including novel/adapted vaccines, long COVID, special populations.
- The relevance of improving knowledge of data sources and their characteristics and how to leverage them during emergencies.
- The importance of preparedness, collaboration and transparency.

Future directions are focused on establishing a framework for RWE generation, including through the VMP, to support continued preparedness and leveraging international RWE guidance.

Olaf Klungel (Universiteit Utrecht, MWP) provided the academia perspective on considerations for the conduct of pharmacoepidemiology studies in the context of the pandemic:

- There was a need for fast generation of evidence. Many non-interventional studies on efficacy/effectiveness during the COVID-19 pandemic were unlikely to provide valuable evidence in the short term, showing that there is a strong need for collaboration between data holders and authorities in order to generate timely evidence.
- Ascertainment of the exposure was key especially in the COVID-19 context in which misclassification and immortal time bias were very likely to affect studies.
- Incomplete information on risk factors was a challenge, as was proper adjustment on these time-varying factors.
- Rapid assessment of the evidence was critical. A key aspect was the transparent reporting using guidelines and publication of reports.

Finally, the importance of reflecting on these lessons learned from the conduct of pharmaco-epidemiological studies during the COVID-19 pandemic and identifying what can be done better in the future was underlined.

Jean-Michel Dogné (PRAC) provided six lessons learned from the COVID-19 pandemic from the PRAC perspective:

- Preparedness is key and the use of harmonised protocol templates for both public and private stakeholders is helpful. Some dedicated guidance for harmonisation will help safety monitoring. The ACCESS project funded by EMA is a very good example, where the data generated have been often used by the PRAC.
- Flexibility and proactivity when new signals are identified, with e.g., need to update background rates by age and gender groups and generate additional background rates.
- Effectiveness and vaccine coverage data are important for benefit-risk assessment.
- Measuring the impact of pharmacovigilance activities through impact research is important to facilitate the implementation of new regulatory measures.
- Collaboration with platforms such as DARWIN EU® and the VMP are valuable to support the work of the PRAC in terms of informing risk minimisation measures.
- During the pandemic, there was an immense amount of data generated as a result of the pharmacovigilance activities. The EudraVigilance team had to adapt the system and increase functionality and extractability to generate evidence in a short period of time. Further work should build on this to improve the way we work.

Nicolas Praet, on behalf of Vaccine Europe, underlined the need of developing robust risk-management plans in a short time. As manufacturers, main challenges in terms of generating and interpreting RWD were:

- Data access and integration – accurate collection of data was not guaranteed due to the specificities of the immunization programmes against COVID 19 which may have occurred outside of traditional vaccination pathways.
- Resource constrains – resources were limited and there was a high demand from stakeholders, creating competition to access/generate data.

- Study design agility – the need for studies that were scalable and able to adapt due to the unpredictability of the pandemic (eg., emergence of variants and need for boosters).

One of the solutions to face these challenges was the cross-industry collaboration, which allowed to coordinate consultations with regulatory authorities, and the establishment of Public Private Partnerships (PPP), such as COVIDRIVE, which allowed to effectively use resources, to quickly generate data covering wide geographical area, to minimise the risk of competition for data access and to align on study designs. Key possible improvements for the future are the need of clearer and globally harmonised guidance from regulators and streamlined protocol review process; improving real time access to data by addressing challenges on data privacy, interoperability and contracting; increasing the use of multi stakeholders' study protocols.

Panel discussion

On the generation of RWE for public health emergencies, key messages were:

- Stakeholders recognised the need for fast access to data. However, data availability is frequently a bottleneck. In this regard, EMA plans through DARWIN EU® and framework contracts with consortia specialised in vaccine surveillance to put in place a system for near real-time data collection and analysis to contribute to active surveillance.
- Data quality is another challenge, as data that could be accessible faster than others might have lower quality. However, it should be recognised that low-quality data available in the short term might still be useful for certain evaluations (e.g., for signal detection).
- Healthcare databases are important source of data for public health emergencies, but they often lack real-time data, which is crucial for decision making. In this regard, the Nordic countries perspective is moving towards having near real-time availability in the near future.
- Another challenge is the diverging results across countries regarding association between vaccines and adverse events.
- Industry recognises the importance of working together, i.e., through multi-stakeholders' approaches, to overcome some of the constrains and contribute efficiently to the efforts of generating RWE.

Closing remarks

The meeting was concluded by Peter Arlett and Jesper Kjær. They thanked all the participants for the rich discussions and productive workshop. It was highlighted that use of RWE regulatory decision making and public health in general is evolving in a very positive way, i.a. with sustainable funding for initiatives (such as DARWIN EU®), with the European Heath Data Space (which will be transformative in the years to come) and with the revised pharma legislation. By collaborating, working together with multi-stakeholder approach, we can be confident that we will enable use and establish the value of RWE by 2025 and beyond; collectively we will contribute towards the excellence in clinical evidence.