



**Fourth virtual meeting of the CIOMS Working Group XIII  
Defining Intent, and Guiding Harmonization and Ethics Standards for  
Real-World Data and Real-World Evidence in Regulatory Decision-Making**

**23-24 September 2020, Virtual Meeting**

## **Minutes**

### **Participants**

Enrica **Alteri** (former EMA), Yoshiko **Atsuta** (Japan Data Center for Hematopoietic Cell Transplantation), Elodie **Aubrun** (Novartis), Laurent **Azoulay** (McGill University), Elodie **Baumfeld Andre** (Roche), Stella **Blackburn** (IQVIA), Monika **da Luz Carvalho Soares** (Agência Nacional de Vigilância Sanitária), Sean **Hennessy** (University of Pennsylvania), Sanna **Hill** (CIOMS), Alar **Irs** (State Agency of Medicines, Estonia), Akihiro **Ishiguro** (Pharmaceuticals and Medical Devices Agency, Japan), Solomon **Iyasu** (Merck, Merck Sharp & Dohme Corp), Michele **Jonsson Funk** (University of North Carolina), Juhaeri **Juhaeri** (Sanofi), Andrea **Machlitt** (Bayer), Miguel-Angel **Mayer** (Universitat Pompeu Fabra Barcelona), Robertino **Mera** (Gilead), Lembit **Rägo** (CIOMS), Anja **Schiel** (Norwegian Medicines Agency), and Kristina **Zint** (Boehringer Ingelheim).

### **Second day only**

Britta **Haenisch** (Bundesinstitut für Arzneimittel und Medizinprodukte), Steffen **Heß** (Bundesinstitut für Arzneimittel und Medizinprodukte), and David **Wormser** (Novartis).

### **Regrets**

Thomas **Brookland** (Roche), Elisa **Gomez-Reino** (Alexion), Lu Hong (National Medical Products Administration, China), and Andreas **Rudkjoebing** (World Medical Association).

### **Alternates did not attend**

Kinue **Nishioka** and Daisaku **Sato** (Pharmaceuticals and Medical Devices Agency, Japan).

## **DAY 1: 23 September 2020**

### **Introduction**

- Lembit **Rägo**, Secretary General, CIOMS, welcomed the WG members and chaired the meeting.
- Lembit welcomed Anja Schiel from the Norwegian Medicines Agency, who is also the Chair of the Scientific Advice Working Party, EMA. Anja has expertise across regulatory and HTA areas.
- Due to the pandemic, the US FDA has taken some time to reflect about joining the WG. CBER has declined to join but CDER may send a representative. [Since the meeting, two FDA contacts have confirmed joining the CIOMS WG XIII.]
- Health Canada may join in time.

- Sanna served as rapporteur.

## Breakout rooms

The WG members were divided into their chosen groups.

Chapter 1	Chapter 2	Chapter 3	Chapter 4
Introduction	Uses of RWE in the regulatory process during the product life-cycle	RWD and data sources	Key scientific considerations in (regulatory) RWE generation
Elisa	Akihiro	Andreas	Andrea
Sean	Alar	Juhaeri	Anja
Stella	Elodie A	Miguel-Angel	Britta
	Elo B A	Robertino	David
	Enrica	Steffen	Laurent
	Hong		Michele
	Kristina		Solomon
	Monica		Yoshiko
	Thomas		
Sanna – only for this day	Lembit – only for this day		

## DAY 2: 24 September 2020

### Chapter teams' presentations

#### Chapter 1. Introduction

- Sean gave the Chapter 1 team's presentation and the notes below reflect only the WG's discussion points.
- The definition by IMI Get Real was put forward as it is among the most explanatory in terms of what is included under RWD and was accepted for being broad and complete enough.
- We need to specify what lies outside the collective conventional e.g. outside interventions and clinical practice, such as:
  - Registries based on epidemiological studies (not clinical practice), which provide the natural history of diseases and information on patterns;
  - Patient-generated data e.g. from wearables or in-home use.
- The source can help also to distinguish what lies outside the standard (outside RCTs) e.g. electronic health records, billing and claims data, product and disease registries, patient-generated data, data gathered from other sources such as mobile devices, social media, and query logs.
- Defining "RCT" and "conventional" would be challenging. Would "case report form"-based be meaningful? Randomised trials use healthcare data and biometric devices also for follow up. Conventional randomised studies do not do those things.
- [Post meeting comment from Enrica regarding whether "case report form"-based could be meaningful: "I think the big divide is the following: trials, even registries, are set up with the specific PURPOSE of collecting prospectively certain health data, usually related to an intervention in a condition. "Real world data" by definition is health care data that is not prospectively collected to investigate a certain intervention, although of course many of them are relevant for that purpose, provided they are understood. Claim data are a case in point. They are not meant to collect efficacy or safety data, but could potentially be relevant for either".]
- FDA or EMA may have a middle-of-the-road definition for "randomised trial". The meaning of the word "conventional" may evolve over time and become unclear.

- RCT definitions will not touch on the types of data involved, as the data itself is not at the core of a RCT.
- The WG can write its own definition for the context of this report, perhaps modifying an existing definition.
- A definition that is too limiting may exclude certain data at a later stage.
- RWE is evidence derived from RWD.
- It may be helpful to provide some historical context. In the past, terms like “evidence-based medicine,” “observational studies,” “observational medical outcome,” and “theory of effectiveness research” were used in this context.
- Other terms to define for use in this report? “Fit for purpose”?
- Explain the relationship/overlap between the terms “epidemiology” (tool, method, science) and “RWD” (source, object).
- The report will have a glossary where we can include the majority of the terms and definitions.
- We will need to clarify the variety and differences of the definitions depending on the scope and objectives of the report.
- We should recognise the definitions from the US FDA, EMA, Health Canada, and PMDA, as they all have their own similar definitions. The regulatory agency of China has an expansive definition of RWD, which includes traditional medicines too under interventions.
- RWD and RWE are increasingly used for studying efficacy, not only safety, and the natural history of disease. [Post meeting comment from Enrica: “I would use in this context the term “effectiveness” instead of efficacy. Efficacy of new interventions/drugs almost by definition cannot be looked at in the real world, because it would not be marketed and utilised...”]

## **Chapter 2. Uses of RWE in the regulatory process during the product life-cycle**

- Elo gave the Chapter 2 team’s presentation and the notes below reflect only the WG’s discussion points.
- We need to describe the importance of the product and the context.
- We need to capture the continuity of evidence generation planning, the proactivity of planning, of generating evidence regardless of where the evidence may come from, which goes beyond the static of pre-authorisation/post-authorisation.
- Before data is generated, at the planning stage, we can be thinking about the quality and value of data to be generated.
- There are novelties that did not exist five to ten years ago: data analysis methodologies and thinking about data in a broader sense, e.g. artificial intelligence and other methodologies. They go beyond the standard epidemiological approaches, which have served well the RWD environment. Different/broader data sources require different analytical capabilities.
- What could be the roles of patients with collecting and using RWD?

## **Chapter 3. RWD and data sources**

- Juhaeri gave Chapter 3 team’s presentation and the notes below reflect only the WG’s discussion points.
- [Post meeting comment: “Apologies for missing this part of the conversation... I think that we should at the very least list other possible data sources, such as biobanks. They are extremely relevant in today’s world, and they beg a good conversation about ethics, and data ownership”.]
- [Post meeting comment: “A very prominent element missing here is the one of data ownership – this has to be tackled, jointly with data sources.”]
- The inclusion of “news” was questioned. Juhaeri explained that news is used, there are tools to search and combine such that machine learning takes information from news, publications and social media. If we disregard quality, accuracy etc, this is part of the current ongoing trend.
- In context of traditional data sources, particularly in relation with healthcare databases and registries, perhaps we should classify items according to physician-reported data and patient-reported data, albeit this is more about information design of prospective cohort studies.

- Some new approaches from the US to cover:
  1. Pre-processed data / ancillary data e.g images: this is a new, different data source that allows for faster data processing and more reproducible analysis. It comes with a data model, and e.g. with response rate for oncology. This is not necessarily for RWE but for health systems research. It gives a semantic layer on top of the common data model. Lembit raised the challenge of a) verification backwards, b) being sure that it is real data and not machine-generated data sets, and c) the possibility of generating things that do not exist. In response, Robertino explained that this is joint work with the FDA. The data comes with e.g. biomarkers and you need to have a common understanding for the outcomes.
  2. Data tokenisation: overcomes issues of confidentiality and privacy, but provides completeness of information for a patient.
- Is the plan to include randomised versus non-randomised data, so not a limitation of the data, but a limitation of the use of RWD to generate RWE? Yes, there is a school of thought that clinical trials can eliminate the potential confounding factors. This is a limitation of RWD, as there may be certain confounders that cannot be controlled. The notion of differentiation between RWD compared to trials (certain things you cannot do in trials) that should be there. Part of it is inherent here in terms of strength in real-life settings in comparison to trials in a confined population.
- Some say randomisation in trials is not necessarily better than accepting confounding in observational studies. We could present the two schools of thought:
  1. True believers of randomisation;
  2. Opposite.
- It is not either/or but more where do we use one and the other.
- The WG agreed on introducing this concept in Chapter 1, and discussing it in-depth under Chapter 3, and in the conclusion.

#### **Chapter 4. Key scientific considerations in (regulatory) RWE generation**

- Michele gave the Chapter 4 team’s presentation and the notes below reflect only the WG’s discussion points.
- Will need close coordination between Chapters 2 and 3 on repeated subjects. Concepts like RWD need to be discussed in a balanced way and not oversold. We need to recognise that there are different, specific intents and that RWD is not for everything.
- Regulatory agencies have become more open to accepting RWD where RCTs cannot be done e.g. with a small population or where there is a lack of evidence. [Post meeting comment: “Lack of evidence of what? Unclear”]
- Under “type of decisions”, include “special populations” (elderly, children, women of child-bearing potential), as this is where we typically glean more data as we have more exposure, and where we would like to have more data.
- Should we include somewhere the concept of “good enough”? E.g. if there is something with a long follow up, the ideal might be a CT, but it may not be feasible for 10-15 years due to drop-outs and high costs. RWD may result in imperfect answers but it may be better than no answer.
  - This could go under trade-offs and sources of bias. No study is perfect.
  - There are conditions where evidence is good enough to support a decision because there is no other way, for ethical reasons or because an RCT is not feasible with a rare disease. Certain considerations drive that decision with limited evidence. Maybe this could be folded into the decision-making chapter, because decision making is about the totality of evidence, and perhaps we can address the weighing of all trade-offs?
  - The decision can be more complex depending on the study. E.g. in a study with an outcome incidence of lung cancer, the intensity of smoking is extremely important, and if this information is missing in the data sets, it could be a no-go. We would need to

consider this when planning the study but also when interpreting the results. There can be critical issues that would stop a study going forward.

- We need to be careful with how we present the “good enough” concept because decision makers, especially regulators, compare the “good enough” with a “substantial evidence” standard. We are not compromising on the evidence standard that we have always used. Even if the best we can come up with under the condition/setting/disease is only case studies (there have been approvals based on case reports), if the evidence is overwhelming enough, even if not based on a traditional RCT, a decision can be made. We need to meet a substantial evidence standard.
- Regarding some parts in the design section, maybe the single arm data will be feasible for e.g. rare diseases, if the numbers are very limited to perform RCTs, and case series, e.g. when it is not ethical to perform large trials.
- Study cost is not to be included in evaluating if something is good enough. We should keep with scientific considerations only.
- RCT and RWD trials should not be allowed to repeat until the results fit what is wanted. To avoid this, typically we need to compare the parity of exposure.
- We could blind the protocol and blind the statistician who is doing the propensity score to the outcomes. This would be a new type of blinding exercise only for this type of exercise. No one has access to the outcome data while we set up the external control arm.
- There is a tension between wanting to be able to lock down information to prevent any maleficence, and if things are not related to the outcome and you put them in your propensity score, this is a disservice to estimating reasonably precise effects. Is there a way to prevent individuals from looking at the outcomes across both groups, but still allow us to confirm our understanding of the patient factors that are strongly related to the risk of the outcome, so that our propensity score models are informed by that? Just throwing a bunch of instruments into the propensity score is going to do bad things too.
- There are algorithms, e.g. the high-dimensional propensity score (hdPS), which use the outcome to empirically select variables that are more likely to be confounders.
- The concern about being able to cherry-pick the model, what is in it, and how you adjust for things, permeates all the observational study considerations. It feels like these are a bigger set of considerations that we should raise.
- Regarding external control arms, we are at a stage where regulators have accepted taking decision in this manner for certain populations, e.g. where we cannot not run a clinical trial, but this may become more commonly accepted if in the future we have the right data for other populations too, methodologically speaking.
- The new concept of “external data” refers to data from RWD sources and other RCTs.
- Interaction between history bias and other bias is important.
- Study design approaches to produce RWE that could support a decision are case specific.
- There were suggestions to add text on:
  - Confounding by time;
  - How to deal with missing data;
  - The importance of sensitivity analysis;
  - Transforming data for reproducibility;
  - How data needs to be fair. Perhaps for inclusion in relation with e.g. tokenised data?
  - Data integrity (i.e. not suspicious or fraudulent data):
    - This is mentioned a little in Chapter 4 under “data sources” in conjunction with data quality and quality control, assurance and audit.
    - There is probably some overlap with Chapter 3. The two chapter teams need to reflect on how the chapters relate to each other, also in terms of chapter length and balance.
    - Perhaps for inclusion under “pre-processed data”? In terms of deciding what role any given construct has in analysis, we may want to favour something that

is of more sensitive or specific construct. If the processing has already been done for us, it is hard to explicitly dial that in for a given research question. This was noted under “Measurement error and misclassification”. Any of the algorithms for identifying these things are going to be imperfect, and knowing how to use this in a given research question, is an important part of the decision.

## Chapter 5. Ethics, governance and related issues

- Data equity;
- Transparency;
- Ethics - In Europe, data privacy is tight. Is it possible to use RWD, who needs to own it, who decides, what are the ethical boundaries (in addition to technical and analytical boundaries)?
  - Lembit is working on bringing on board specialised ethicists with the appropriate legal backgrounds.
  - Laurent proposed the ethicist, Jonathan Kimmelman, who has published on trials and RWD. <https://www.mcgill.ca/biomedicalethicsunit/faculty/kimmelman>
  - [Post meeting comment from Enrica: I do not think we should mix ethics with data privacy, which in EU is regulated very strictly. The impact of GDPR on any use of RWD is important to discuss. I attach the link for reference. <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679&from=EN> However, ethical issues are not solely related to data privacy regulations, and data ownership. It may be unethical NOT to use health data to foster research in rare diseases. The question is who decides? For what purpose? I suggest that we separate data privacy and compliance with GDPR from ethical issues, where patients’ representatives may bring a good perspective to the WG.”]
- Data governance – We need to have good governance for data. There is a lot of divergence on this matter and it continues to evolve.
  - How organisations/companies are set up is a challenge: who owns the data and who does what? There exist cultural differences.
  - Organisational approaches within regulatory agencies can vary, such as: processes, submission, engagement, and consultation. Organisations and regulatory agencies may not be set up in the right way and have the right people to optimise the use of data for decision making. How do we re-organise to optimise the generation of RWE, and how do we make decisions about whether to go forward with a RWE activity or not?

## Next steps

- Start drafting. Each chapter team to decide how they prefer to organise their work.
- The next full WG meeting is to be scheduled for mid-December 2020.

## Any other business

- CIOMS reports have been fairly unique and of varying length in the past, ranging between 100-200 pages in A5-size booklet format. Length will depend on the writing style, appendixes, and editing in the end.
- The full report is to be completed in approximately two years, with first draft delivered by early December 2020.
- Does the report need to be endorsed by member organisations? No, the CIOMS WG will endorse as we consider the members to be experts; it is a consensus report. The regulators and perhaps others too, will typically require internal acceptance.
- Towards the end of the writing process, we will set up a smaller editorial team but the final outcome will be reported the full WG.
- Case studies can be placed in the chapters or in appendixes depending on their length. The longer ones will fit best in the appendix.

- CIOMS produces print copies and issues electronic versions of its reports. Some of the contents may have rapidly-changing elements and these can be hosted online. This was done for the recently published CIOMS publication, the [Drug-Induced Liver Injury](#) (DILI) report. The CIOMS WG XIII can decide.
- CIOMS can assist with organising the chapter meetings.