Council for International Organizations of Medical Sciences

9th meeting of the CIOMS Working Group WG XIV on Artificial Intelligence in Pharmacovigilance

24-25 September 2024, Darmstadt, Germany, hybrid meeting

Meeting Minutes

Overview:

Summary

The CIOMS Working Group (WG) XIV on Artificial Intelligence (AI) in Pharmacovigilance (PV) held its 9th meeting in Darmstadt, Germany, on 24-25 September 2024, with participants attending both in-person and virtually. At the 9th meeting, the subgroups worked on the chapters – the report had been re-structured according to common principles – and the use cases; and the WG looked ahead to future milestones for maturing the draft e.g. a Public Consultation. The next in-person meeting will be held in Geneva, Switzerland, in late April or early May 2025, while subgroups will continue work virtually.

Minutes of discussion

Day 1

1. Opening and welcome

- Hervé le Louët, CIOMS President, welcomed the members to the 9th meeting of the CIOMS WG XIV on AI in PV. He congratulated the members on the progress that has been made during the past months and reminded everyone that CIOMS is a forum for free thinking and innovation.
- Lembit Rägo, CIOMS Secretary General, added his words of welcome and thanks. He opened the meeting as chairman for the two days (for a list of participants see Annex 1).
- Lembit made the following announcements:
  - The Introduction to MedDRA Labeling Grouping (MLG) report was published electronically in March 2024.
  - The CIOMS WG XIII's report on Real-world data and real-world evidence in regulatory decision making was published in May 2024.
  - The CIOMS Cumulative Glossary with a Focus on Pharmacovigilance – 75th Anniversary Edition was published in September 2024. This combines all the WGs' terms and definitions into one product and is a very good reflection of CIOMS's Jubilee celebration.
  - CIOMS has also produced a glossary of ICH terms and definitions as part of our more structured collaboration with The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH). We are in the process of producing a memorandum of understanding with ICH.
  - The WG XII report on Benefit-Risk Balance for Medicinal Products is at a mature stage and may reach publication before the end of 2024.
  - The WG report on Severe Cutaneous Adverse Reactions – SCAR is also at an advanced stage. This report is the latest in a series of organ-specific guidance documents, where the first was a report on Drug-Induced Liver Injury, and we hope there will be more in this series.
- Regarding the agenda, there was a request to discuss large language models (LLMs). Do they fit under general principles or do we need stricter principles? What would be unique about LLMs?
- The CIOMS support will be handed over from Kateriina to Sanna. This is to bring more time availability for editing and experience with e.g. the Public Consultation process.
- Hans-Jörg provided logistics for the days and historical details about Merck and the building.
- Sanna was rapporteur.

2. Reflections on the Artificial Intelligence landscape and recent developments

Tour de table

- The European Medicines Agency published its guiding principles: Harnessing AI in medicines regulation: use of large language models (LLMs).
- In the context of a critical appraisal of literature, Benny mentioned that AI alone is not enough for critical thinking, but if it is used as a tool/assistant, it can bring value. This may be useful in a case study.

- The European Union's AI Act touched on high-risk AI.
- The EMA's Reflection paper on the use of Artificial Intelligence (AI) in the medicinal product lifecycle is due to be published in the coming weeks.
- The FDA's CDER Emerging Drug Safety Technology Program (EDSTP) is under way to garner industry perspective.
- Fairness and equity are important topics especially in countries where AI regulation has not been applied yet and where there exist high social inequity and where there is a huge lack of data.
- Since the CIOMS WG XIV was established, there has been a continuous change in the landscape and e.g. deeper integration of AI in devices. We are using AI that has been developed for other purposes and that changes our thinking. We have our models, situations, use cases, and data sets; and a lot of the principles do still apply in these circumstances, but how we apply them may be slightly different. Consequently, we may need to re-frame some of the language in the report.
- Over the past half year, every facet within a company has been asking how they can leverage AI but it is not necessarily a magic bullet. There are uses of AI on the fringes of PV, e.g. document management, training using large data sets, and large document summarisations; but it is challenging to use AI to aggregate data and analyse it for PV.
- The WG XIV's principles would be expected to stand the test of time even if the technology changes.
- On the 5th of November 2024, the annual HMA/EMA multi-stakeholder workshop on artificial intelligence (AI) - enabling the safe and responsible use of AI will take place. Follow the link for practical information including how to register. The event will showcase a number of use cases and descriptions.
- From the WG XIV's 1st meeting to the 9th meeting, the landscape has evolved, and we would like the report to stand the test of time. AI is becoming a part of how we work. The limitations of how we bring in AI from a PV perspective is another emerging challenge.
- Denny was pleased with how the WG has expanded on the principles and put the grid (of principles) into place.
- The Chapter 13 on Future Vision has not been updated during the last six months. We would welcome at least one regulator to review it. Some of this chapter may have come into practice and is not to do with the future any longer, and it may need updating with new visionary aspects.
- While the CIOMS WG XIV report has been maturing, other initiatives have been progressing too, e.g. various government documents and the output from the International Society for Pharmaceutical Engineering (ISPE) including the GAMP 5 Guide 2nd Edition with a chapter on how to validate an AI-built system. This is worth reviewing. Furthermore, a Good Practice Guide by ISPE for handling AI systems in GxP environment is on the way.
- The number of professionals with the technical knowhow for developing AI systems is growing but the specialists in PV is not growing at the same pace. The CIOMS WG XIV report perhaps needs to be more PV-focused. Some examples may help in this regard.
- After publication, it will be interesting to see the promotional activity that may follow from e.g. CIOMS, EMA, and some UN organizations.
- Scalability is still an issue and there is some reluctance to adopt AI. The reluctance may be due to unclear regulation and guidance, concern about risk taking, and data privacy as a lot of solutions are cloud-hosted.
- The Pharmaceutical Inspection Co-operation Scheme (PIC/S) AI and ML in PV group has chosen November 2025 for its deliverable. The scope is currently undecided but the UK's Medicines and Healthcare products Regulatory Agency (MHRA) has offered to host the event. The CIOMS WG XIV report timeline being unclear is a concern.

- Japan's Pharmaceuticals and Medical Devices Agency (PMDA) has not published guidance on AI in PV but it continues to discuss and communicate with industry and academic groups to improve its understanding of new technologies.
- How do we go from a pilot concept to putting it into production? What are the validation frameworks? What are the risks we have to take? Should we have one or two additional use cases in the report to show what LLMs look like?
- In the context of infectious diseases, there is a lot of epidemiology data generated from AI tools and there seems to be a gap in the AI tools being used in epidemiology. This is related to work in PV when we write protocols, what we do with the post-authorisation safety study (PASS), what we do in any kind of epidemiological study, the risk management plans (RMPs), and how much of that can be AI-generated. That is important because it is about time as well as quality. There are tools that may be very useful for PV and for generating e.g. Periodic Safety Update Reports (PSURs), RMPs, etc.
- How do you make AI systems GxP-compliant? How do we scale up expertise? There are more and more training programmes being run for Generative Pre-trained Transformers (GPTs).
- The Foundation for the National Institutes of Health, among its other programs, focuses on enhancing vaccine safety surveillance including via PASSs. In low- and middle-income countries, there is a potential application of AI to enhance ability to better understand vaccine safety. Contact Walter for more information and details for who to contact at NIH, FDA and Johns Hopkins University.
- There is a new/enforced awareness about how important it is to identify the right LLM for a specific task and use the right set of data for the region in question.
- The AI discussion has progressed slowly, e.g. looking at hallucinations: a year ago, hallucinations were taken to mean that something was being done wrong; then, we progressed to acknowledging the existence of hallucinations and how to prevent this happening; and now, we are questioning under what circumstances and what systems around a LLM do we need to prevent problematic hallucinations from occurring in a PV system.
- Issues that are foundational to PV are critically difficult questions for AI, and as long as we focus on the problems in PV, our challenges and opportunities, how do we use AI to further those, then we are going to be future-proofed around the capabilities.

3. Progress updates on chapter drafts and use cases

General reflections

- The report being re-organised around the principles seems to provide better clarity.
- The structure and content are improved from the last version.
- There remains work to do on the Introduction, Landscape, Scope and the Future Vision chapters.
- There were discussions around the term Principle / Definition. Are these really Definitions? Should we call them Principles? Another term used was Description.
- There was a suggestion for re-organising the structure a little further (TBC):
    - Introduction
    - Landscape
    - Scope
    - Create a summary of Lifecycle Phases - lifecycle varies across the report, whether it is around development or pre- / post-implementation. We would need to define lifecycle.
    - Principles

  - Move Governance and Accountability as last principle
  - Consider addition of "Human Workforce considerations" chapter (more to follow)
  - Future State
  - Appendices: use cases, glossary and references
- Introduction
  - There is a need to add references where examples have been made to illustrate points. Ideally, when possible, we use PV examples, but when they are not available, we provide other examples to illustrate points.
  - We need to add a statement to say that the report applies to all AI, including generative AI.
- Table of principles.
  - The descriptions were developed for the purposes of this document.
  - We are not trying to override definitions that have been provided by health authorities.
  - We include a statement that it is important to review these principles in the context of an application for a specific situation but not all principles will be relevant for all scenarios.
- We need examples to reinforce points. Ideally, publish citations as much as possible, but we also recognise that there has not been a lot published in the space.
- Landscape analysis.
  - This was not covered at the 9th meeting but it was acknowledged that landscape analysis with the publication analysis will need to be revisited.
  - With each of the chapters, when they are citing an example, they are trying to emphasise a specific point, whereas the landscape may be a bit broader.
  - Much of this work was done by Niklas and there are some additional references from Robert to be integrated. The effort needs to be shared out.
- Human workforce.
  - Various chapters have related content, e.g. Fairness and Equity, and Human Oversight, which could be combined into a separate, specific topic about "humans".
  - This could include subjects such as human-in-the-loop to "human" in general.
  - Other topics could be included, e.g. change management and organisational readiness, depending on how the chapter evolves.
- Quality assurance.
  - Robert and Taxiarchis are carrying out a literature review on quality assurance for AI and PV, and the informal conclusion at this stage is that there is very little available.
  - There is not much practical advice on what can be done to make sure that AI tools are working in practice.
  - The review is about how quality assurance is done in PV and how quality assurance is done in other disciplines for AI applications.

- The following four priority chapters were discussed on Day 1:

Chapter 12 on Fairness and Equity

- Beth presented the Fairness and Equity chapter team's work.

- We recommend a high-level, general summary about ethics and Fairness & Equity, and how these two subject areas tie together as there seems to be some overlap. This could be early in the report, in the Introduction, or potentially within the Description of Chapter four.
- We need to define what is meant by fairness and what is meant by equity. These need to be included in the glossary.

Chapter 8 on Validity and Robustness

- Niklas presented the Validity and Robustness chapter team's work.
- The chapter team spent time on the Definition/Principle, which was edited further, and on the key messages.
- Regarding the Definition:
  - "Validation and Robustness" has been edited to "Validity and Robustness".
  - On the first bullet, small edits were made: "*Validity means that a system achieves its purpose within acceptable parameters. It requires predefining acceptable performance parameters, selection of data for model training and/or testing …*", we recognise that we do not always have training with zero shot learning etc, so we want to make sure that we do not over specify. We can have either training or testing, and then assessment and model performance in a realistic setting and deployment into an ongoing quality assessment process.
  - On the second bullet: "*Robustness means that a system reliably performs its intended objectives accounting for variations in data*", we believe the meaning is a little improved from editing and prefer to retain the second bullet, despite a suggestion to delete it, feeling that it adds value. This can be discussed further if needed.
- Regarding the Key messages:
  - Minor edits were made to the key messages to help clarify the text.
  - A fourth broader key message was added (drawn from the chapter): "*There should be sufficient representation of relevant types of data in the evaluation set to detect biases, promote adequate and generalizable performance across the intended deployment domain, assess usability, and identify circumstances where the model may underperform.*"
  - Robert suggested integrating possible solutions into the bullet rather than focusing on problems with evaluations. Some years ago, the FDA published a paper on AI in PV that mentioned alternative models or approaches. Robert also related an example of Stevens-Johnson syndrome (SJS), where the model for SJS alone would not be helpful, but where multiple models running in parallel may help to ensure quality.
  - How do we build large enough and representative enough data sets? The new point is more technical and may require some content in the chapter on methods for creating a large diversion of data sets.
- There was a discussion about computer systems validation and the inadequacy of some of the standard computer system validation processes for testing or validating general computer systems when we have an AI solution or another solution that changes rapidly over time.
  - In the chapter today, we say what we should be doing.
  - We do not say that some of the normal things do not apply in context.
  - We actually say that we should also be doing computer systems validation.
  - Certain inadequacies exist and currently our text does not acknowledge that AI in PV does not always work. We need to draft new text.

- There are overlaps/connections between the Fairness and Equity chapter and the Validity and Robustness chapter, e.g. an algorithm may perform well globally but we get less good subgroup performance evaluations, e.g. in a specific part of the world or for paediatric groups. We will need to strengthen connecting back to the Fairness and Equity chapter.

Chapter 5 on Governance and Accountability

- Denny presented the Governance and Accountability chapter team's work.
- To make the report practical, we will list the principles in a grid and consider the medicine development lifecycle and practical examples, both published and unpublished cases.
- The grid gives considerations for AI use cases, how to govern them, making them GxP-compliant, in addition to maybe further validation requirements. Should explainability be in the grid?
- How to handle the requirements phase? We have decided to split the requirements phase off. We will zoom into the requirements phase separately and explain what it is and how it refers to the principles, so that we take into consideration all the non-functional requirements, etc. We will elaborate more on the requirements phase in a later draft of the report.
- Chapter teams are invited to review the use cases again as the field is evolving fast in order to make sure that e.g. the risk-based approach is still valid.
- This chapter is complex because it has implications on all the other chapters. All of the principles are mapped out – you can link to the use cases that adhere to the principles.
- All of our organisations are different and that makes being prescriptive about government, even governance principles, difficult.
- It might be difficult for a very small organisation to adhere to the principles.
- Reflections on the placement of the governance chapter:
  - We get into governance before we explain what we are governing;
  - Readers will dip in and out of the report, and read the chapter(s) they are interested in;
  - The closely related chapters are in a particular order for a natural flow;
  - Some overlap between sections is to be expected.
- The grid was moved to the end of the document as a summary, putting the guidance into practice.
  - Some of the comments were addressed and edits were made directly into the GoogleDocs file.
  - The chapter teams are invited to review and comment on the grid key words.
  - We addressed how different types of organisations can benefit from governance and accountability.
  - We removed the sentence saying that accountability for AI systems cannot be outsourced. The Qualified Person for Pharmacovigilance (QPPV)-type activities can be outsourced but accountability itself cannot be outsourced. There is a new section on this.

Chapter 6 on Risk-Based Approach

- Julie presented the Risk-Based Approach chapter team's work.
- There were some minor edits on the Definition (Principle).
- The team is considering making a reference to the CIOMS WG IX definition of Risk and will make a decision about this at a later stage.
- We have added one key message on how PV processes need to take into account that AI algorithms are not perfect.

- In the regulatory considerations section, we have acknowledged the challenges around the dynamic nature of regulatory requirements in the context of the fast-moving AI technology, and this will be further discussed in the Future Vision chapter.
- We have acknowledged that human-in-the-loop might not be the only or the best solution in some instances.
- The issue detection and mitigation section has been extended in scope to also cover the development phase. We were putting more emphasis on the post-deployment monitoring activities, and so we have adjusted the content on this point.

DAY 2

4. Progress updates on the chapter drafts and use cases - continued

Use cases group

- Arvind presented the use case team's work.
- The use cases were reviewed against the updated principles definition and the new information available.
- It is important that the way we describe the principles on the grid is the same way as they are described in the use cases. We need ensure we are aligned on the governance structure, the framework, and the principles.
- The team identified five actions:
  1) Add a preamble to the use case section to give context, explaining to the reader that this is more a retrospective review of the principles against the use cases.
  2) Where applicable, identify and add potential future opportunities. Every use case provides a status and we recognise that some opportunities continue to improve. For example, in the translation use case, the opportunity is to shift from 100% quality control to a sampling approach.
  3) Re-consider how to display the summary table. The colour-coded table was intended to be a visual overview approach to viewing the report. It identifies the different stages of alignment to the principles: green for fully aligned, blue for partially aligned, red for not aligned, and gray for not applicable.

     Disadvantages identified with the table:
     - we would prefer to not make a judgement over someone else's compliance;
     - we do not define what we mean by the different colours.

     Maybe we can say that the colours simply suggest a degree of alignment based on the information available? Perhaps we can display the table in grayscale to enable readers to reproduce/print/photocopy it easier?
  4) In the table below the colour coding, where we have the description of the principles and how the use cases meet those principles, we invite the use case leads to review the table for accuracy. We also ask for feedback on if the messages match for each chapter. When we write key messages, we need to rethink about their applicability rather than simply keeping them high level. Do they fit with what we have done?
  5) We have about six use cases, and maybe with a possibility of adding an additional one or two use cases specific to LLMs. Is the WG in agreement?

- The following further four chapters of secondary priority were discussed on Day 2:

Chapter 7 on Human Oversight

- Julie presented the Human Oversight chapter team's work.
- We have addressed most of the comments that necessitated slight adjustments to the Definition (Principle).
- In the key messages, we have nuanced the wording around the role of the Human Oversight.
- We need to expand the section entitled *Types of human oversight*.
- When we introduce terminology, including human-in-the-loop and human-in-command, we need to define what is meant by each of these terms.
- There is an open question: could AI-assisted performance monitoring be considered in/under-validation? This is not within the scope of the Human Oversight chapter. Would the Validity and Robustness chapter like to address this?
- Do we need a standalone chapter to discuss the professional and organisational transformations linked to the implementation of AI systems in PV processes?
- In the current chapter on Human Oversight, there is a section that discusses the transformation of traditional roles. The discussion that followed touched on topics such as:
    - The use of AI for monitoring, i.e. in the context of AI- or ML-assisted human oversight. Would this be suitable during validation activity? Does this fit with the Validity and Robustness chapter or with different phases of the lifecycle or the Governance chapter content? Perhaps this topic belongs across several chapters?
    - Do we mean to suggest using an ensemble of models, i.e. two or more models: AI-assisted monitoring and using a simple rules-based model, plus maybe a legacy system to potentially help us to rely on the output better, although it may still not be perfect? We still have to qualify the monitoring with human-in-the-loop. An ensemble of models cannot replace human-in-the-loop.

Chapter 11 on Data Privacy

- Walter presented the Data Privacy chapter team's work.
- The chapter team is anticipating further comments from the FDA on Health Insurance Portability and Accountability Act (HIPAA) and ethical principles on conducting research on human populations.
- The preamble in the Data Privacy section on ethics that includes the Belmont principles, the US Common Rule, and the Declaration of Helsinki will be moved to the Introductory section.
- We will invite colleagues at the EMA to review the section on the European Union's General Data Protection Regulation (GDPR) for accuracy and completeness.
- The CIOMS WG XIII report on Real-world data and real-world evidence in regulatory decision making Chapter 4 on *Ethics and governance* contains similar content on ethics, data privacy and the GDPR.
- The chapter team would welcome similar content from the PMDA, Japan, ANVISA, Brazil, Health Canada, the Danish Medicines Agency, and any other data privacy regulations.
- Data privacy regulations have been part of PV for a long time. What is changing now is the computational power that allows us potentially to re-identify individuals.

- The team discussed the challenges of LLMs and especially open models vs closed models, and the risk of re-identification, acknowledging that prior approaches to de-identification may be insufficient in the area of generative AI when one can identify individuals based on very limited data points. This has been known for some time but it has become more relevant with generative AI. Thinking about controls, it is possible that, despite best intentions, individuals can add additional data that increases the risk of re-identification. Genomic data carries further risk of re-identification. We also talked about controls. Must take care about how we define open and closed models.
- Synthetic data:
  - The use of synthetic data may be an option to ensure data privacy - there may be scenarios where it may be leveraged.
  - Synthetic data may be mentioned in the Future Vision chapter.
  - This discussion originates rather from public health.
  - The Belmont principles and the Declaration of Helsinki were designed with the mindset of thinking about clinical research, not public health.
  - There are exemptions around the use of data that would otherwise be what is referred to in the US as Public Health Information (PHI), which allows for the use of data that might otherwise be protected for the purposes of public health. There are similar exemptions within the GDPR. It may be possible to reference the CIOMS WG XIII report Chapter 4.
- There were some examples provided around data breaches, which highlighted the risks even in the absence of generative AI. The risks are around the protection of content.
- Data owners can give wide consent if data are under good management practices including wide consent for secondary use of data. The data source is important. Some aspects are elaborated in other documents and information can be referenced.

Chapter 9 on Transparency

- Niklas presented the Transparency chapter team's work.
- There were few changes to the Definition (Principle):
  - Development and evaluation should be covered under Transparency but training data sets are not always involved and so should not be mentioned. Edits were made accordingly.
  - In terms of the stakeholders, the focus is going to be between e.g. developers, PV business owners, inspectors and regulators. The public is relevant but less important in this context.
- There were some edits made to the key messages:
  - Key message 1: "*Declaring when and how AI solutions are used in pharmacovigilance processes is critical for building trust among domain experts, decision-makers, regulatory authorities, and the public.*" This does not refer to normal work tasks, e.g. the use of copilots to write a text, but rather the use of AI when augmenting or automating PV core tasks. We need to clarify the scope upfront, perhaps earlier than the Transparency chapter.
  - We want to add another key message to say that transparency means sharing information between parties and that the type of information that is appropriate to share will vary and needs to be tailored to the recipients.
  - The last key message is probably too narrow: "*Performance evaluation results should combine quantitative metrics and qualitative examples to give a complete picture of the AI solution's effectiveness and limitations for informed decision-making.*". This needs to be revisited, bringing in broader content from the chapter.

- Towards the end of the introduction, a paragraph was deleted about disclosing the use of AI to process and share data intended for PV systems, and for improving AI solutions. We should inform people if we intend to use their data for the development of downstream AI solutions.
- Development and deployment.
    - To what extent do we need to have traceability during method development, and especially at the early phases of method development? We have a trade-off between traceability and efficiency and speed. What about during development?
    - Once we deploy something, and under continued deployment, we need to have very good traceability: we need to know what that version of that model we have, and we need to be able to trace back and make changes to that model.
    - There are different forms of transparency: we may need to be transparent to an inspector and we may want to have transparency for efficiency.
- Under continual integration and deployment, the team added a sentence: "*Transparency between the development team and the pharmacovigilance organization is crucial to ensure efficiency and that the solution is fit-for-purpose*".
- The team also strengthened the language around having good documentation of the development changes when working with a vendor.
- When working with an external vendor, often the question of intellectual property arises, e.g. not being able to share data. Do we have some leeway or do we expect complete transparency? Some will say they cannot be transparent, they cannot say what kind of model they are using, and what kind of input variables they are considering. Then it will become a negotiation: can I use that if you cannot tell me this? It may be that with respect to intellectual property rights, the parties involved need to come to an understanding over transparency. Perhaps a vendor will be transparent with a regulator even if not with other parties? If not, then we still have to prove that our solution is fit for purpose even if it is not transparent and we do not have all the information. The chapter team may write about the considerations, e.g. maximising transparency and the challenges.
- The WG members related some of their experiences of involving vendor contracts.

Chapter 10 on Explainability

- Satoko presented the Explainability chapter team's work.
- The principle wording was edited as "the public" was considered inappropriate and was updated to "the individuals".
- The key messages have been edited and the last bullet concerning the stakeholders is still under discussion.
    - The team softened the language around the need to have complete explainability of how the system arrives at its output.
    - More work needs to be done on human factors such as overtrust – this need to be considered especially in the context of limits and risks, keeping in mind that the interpretation of the explanation could be different from person to person.
- The content under the section entitled *Explainability vs interpretability and transparency* will be rearranged into shorter, concise messages; possibly into a table.
- We will clarify that the definitions – explainability, interpretability, transparency – are for the purpose of this report. These may be moved to the Introduction, giving concise examples of each.

- There is a point about explainability regarding identifying duplicates that needs clarifying in the report. This is an interesting case where AI is actually doing a better job than a human in understanding the statistical unlikelihood of matching a large number of drugs. We did not fully appreciate how strong evidence that was for duplication. In this case, because that model is explainable, we can see what it is doing. The reason it is picking this up is because it is matching six drugs that are not used commonly together. We did not fully appreciate how unlikely that was to happen. Typically, that would not be a data point that we would necessarily weigh heavily if we were doing the matching manually. We would not even look at it.

  The sentence in question is this one: "*Due to missing outcomes, onset dates and ages that were close but not matching, and no matches between the registered adverse drug reaction terms, these cases would not have been identified as duplicates if they were not flagged by the AI. The AI explanation revealed that the match score was based on six different drug substances which were identical between the cases in addition to the fact that these six drug substances are commonly not co-reported.*"

- Under "Limits and risks of explainability", under the subtitle "Risks", some of the bullets may need to be edited or deleted. The bullets highlighted in yellow will be kept. These are about how explanations could be interpreted differently by people or people could overtrust the system. We would also re-write the risks and limitations.
- Explainable Artificial Intelligence (XAI).
  - We could go deeper into these methods or remove some of the detail and keep it more high level.
  - Some of it may even fit into the future landscape.
  - The risk is that it is very technical and it will likely be outdated quickly.
  - There is an ongoing scientific discussion about the validity of these approaches.

5. Milestones for maturing the draft

- The Public Consultation (PC) can enhance the quality of the report.
- Being mindful of time, the most time-efficient approach to carrying out a PC would probably be the following:
  - Prepare the draft report for the PC through virtual meetings by the end of December;
  - The WG members review and feedback on the draft report during three weeks in January;
  - The Chapter leads integrate the feedback over three weeks in January/February;
  - The draft report is posted for PC for six weeks over March and April;
  - The PC comments will be cleaned minimally and inserted into the draft;
  - The WG members meet in-person to discuss the PC comments in late April or May. Discussion can take place in chapter teams and at the WG-level.
  - The Editorial Group can do a final polish before publication.
- There is a fair amount to do to reach a near-final version of the draft report given that the Introduction, Scope and Landscape have not been drafted.
- Some of the comments made included:
  - More work gets done when working in-person as there are less distractions.
  - We can try to address as many comments as possible in advance of the in-person meeting e.g. minor editorial edits and some comments can be resolved by the individual authors, reducing the number of comments that the WG needs to discuss.

  - If all the comments can be resolved virtually in advance of the in-person meeting, we may not need the in-person meeting.
  - It is possible to hold the WG members' review and the PC in parallel, but this was not the preferred approach.
- The purpose of the PC is to identify omissions, mistakes and imbalances in the report. We will aim to get information from targeted companies and organisations including AI vendors. We are not looking for wordsmithing or for commentators who want to us to add their publications into our references, which is the kind of comments we sometimes receive.
- The WG were shown a sample CIOMS PC form used for eliciting comments, showing where commentators fill in their own details and give consent for being included in the final report (or not, as not everyone wishes to be named). The form has a place for general comments, which we will appreciate but may assign a different priority when working through the comments in the spirit of time-efficiency. The form gives an example comment. The WG XIV's comment form will be approved by the WG members before it is uploaded online.
  - All WG member organisations will be invited help raise awareness about the PC opportunity e.g. through social media channels.
  - Once the time for sending in comments has closed, Sanna will insert all the comments into the draft report (Word doc) so that they appear in the correct context.
  - In a transparent manner, the CIOMS Secretariat records a response to every comment so that we can respond to any queries.
- There is concern over version control.
  - It would be helpful to have the draft report in one place (editing platform?) with all the chapters combined into one document for the benefit of e.g. cross referencing within the report. This would make it easier also for the use cases.
  - The Editorial Group has the draft report in SharePoint and Beth has given access to all WG members. On SharePoint, the various chapters are saved separately. It is important for everyone to respect the cut-off dates regarding providing comments and to feedback in the correct place.
- References
  - WG members are welcome to use their preferred reference management software.
  - Please provide as much detail as possible and the CIOMS Secretariat can help with completing and formatting the references in preparation for the PC.
  - [Post-meeting comment to give some help if needed: CIOMS reports use superscripted numerals in the body of the chapter to cite references. We use the Vancouver referencing style. To help the reader, we show the URL where the reference can be found and add the date we last used the URL: '[Accessed 2 December 2020]'. Include hyperlinks to the full paper/PDF/web page (if open-access without login), or else PubMed entries, as shown in the examples below. In this way readers can see whether the full paper is available. We provide the doi number wherever possible.

    - Maxmen A. Busting the billion-dollar myth: how to slash the cost of drug development. Nature. 2016;536(7617):388–390. (Journal full text) https://doi.org/10.1038/536388a
    - Maïga D, Akanmori BD, Chocarro L. Regulatory oversight of clinical trials in Africa: progress over the past 5 years. Vaccine. 2009;27(52):7249-7252. (PubMed) https://doi.org/10.1016/j.vaccine.2009.08.113

6. Glossary

- The CIOMS WG XIV report will have a glossary.
- Douglas accepted the role for leading the glossary.
- A definition of Artificial Intelligence was put forward:

> *artificial intelligence n. a branch of computer science that involves the ability of a machine, typically a computer, to emulate specific aspects of human behaviour and to deal with tasks that are normally regarded as primarily proceeding from human cerebral activity. [Latin: artificialis, from ars, artis, art, and facere to make; intelligens, prp of intellegere, to understand, from inter, between, and legere, to choose]*
>
> Source: Aronson, J.K. Artificial Intelligence in Pharmacovigilance: An Introduction to Terms, Concepts, Applications, and Limitations. Drug Safety 45, 407–418 (2022). https://doi.org/10.1007/s40264-022-01156-5

  - This definition will be adopted into use in the report if no WG member expresses reservations within three weeks following the 9th WG meeting, i.e. by the 17th of October. [Post-meeting comment: as this deadline was not expressed clearly, it has been extended to the 31st of October.]
  - AI is not just restricted to ML. It can include simple, rule-based methods, performing or supportive tasks that would otherwise require some sort of human expertise; it can be a very narrow task. This is a comprehensive definition. For reference, Deep Blue was a chess playing computer system in the 1990s, which did not use ML but embedded human knowledge in the form of pre-defined rules and strategies (and efficient search of possible game trajectories).
  - We probably need to discuss the AI definition in the Introduction because a lot of the above applies or does not apply to learning data. Some of the issues are eliminated when a system is rules based.
  - The WG XIV report glossary could modify this AI definition and acknowledge the original source. We would present it as: "Proposed by CIOMS Working Group XIV. Modified from: Aronson, J.K. ….".
- With respect to the other definitions needed for the report, the WG is welcome to select pre-existing, established definitions, or if ones do not exist that fit the content / messaging in the report, it is fine to create new definitions for the purpose of the report.
- Sanna is happy to review the draft and suggest which terms may be good to define.

7. Any other business

- When WG members have an opportunity to represent the WG at events, it is good practice to inform everyone and share PowerPoint slides in advance for everyone's visibility, an opportunity to contribute content, and ensure that everyone benefits from sharing resources. We can store slides on the WG's online password-protected pages. Currently, there is no disclaimer in use for the WG's work being work-in-progress.

8. Next steps / next meeting

- The next, in-person WG meeting will take place in Geneva.
- The dates will be researched and a Doodle poll will be circulated asap.

9. Closing remarks

Hervé and Lembit thanked the WG members for joining in-person and virtually, and for the productive discussions.

10. Annex 1: List of participants

Attending in person

Justyna Amelio (AbbVie), Andrew Bate (GlaxoSmithKline), Arvind Bellur (CSL Behring), Adrian Berridge (Takeda), Hua Carroll (Biogen), Douglas Domalik (AstraZeneca), Julie Durand (EMA), Kirsten Egebjerg Juul (Danish Medicines Agency), Piero Francesco Franco (Pfizer), Neal Grabowski (Sanofi), Stephen Heaton* (CIOMS), Thomas Henn (Unither), Sanna Hill (CIOMS), Satoko Hirokawa (Health and Youth Care Inspectorate, Netherlands), Dieter Kempf (Genentech, Roche), Hervé Le Louët (CIOMS), Benny Ling (Health Canada), Denny Lorenz (Bayer), Monica da Luz Carvalho Soares (ANVISA, Brazil), Elizabeth MacEntee Pileggi (Johnson & Johnson), Yusuke Matsunuga (PMDA, Japan), Manuela Messelhäuser (PEI), Niklas Norén (WHO/UMC), Lembit Rägo (CIOMS), Kateriina Rannula (CIOMS), Hans-Jörg Römming (Merck Group), Walter Straus (Moderna), and Phil Tregunno (MHRA, UK).

Attending virtually

Robert Ball (US FDA), Taxiarchis Botsis (Johns Hopkins University School of Medicine, US), Vijay Kara* (GlaxoSmithKline), Flávia Moreira Cruz (ANVISA, Brazil), and Panos Tsintis (CIOMS Senior Adviser).

*Alternate

Apologies

Russ Altman (Stanford University, US), Luisa Barrios (Merck Sharp & Dohme), Brian Buch (MHRA, UK), Yauhenia Cherkas* (Johnson & Johnson), Shaun Comfort* (Genentech, Roche), Selin Cooper* (AbbVie), Jean-Michel Dogné (University of Namur, Belgium), Julie Girod* (Sanofi), Kendal Harrison (MHRA, UK), Kostadinos Kidos (Takeda), Yuki Kikuchi* (PMDA), Roli Mathur (ICMR), Richard McAteer* (Health Canada), Eva-Lisa Meldau* (WHO/UMC), Dirk Mentzer (PEI), Ravi Patel (Unither), John Reinhard Pietzsch* (Bayer), Stephen Rosenfeld (North Star Review Board), Irene Scholz (Swissmedic), Thomas Stammschulte (Swissmedic), James Whitehead* (AstraZeneca), and Brian Yau (WHO).