



Second meeting of the CIOMS Working Group XII:
Benefit-Risk Balance for Medicinal Products – Update of CIOMS IV
29-30 April 2020, virtual meeting, hosted by Pfizer

Meeting Minutes

Participants

Patrick Caubel (Pfizer), Guacira Corrêa de Matos (Anvisa), Scott Evans (GWSPH), Stephen Evans (LSHTM), Richard Forshee (US FDA CBER), Stewart Geary (Eisai), Takahiro Goto (WHO), Luther Gwaza (WHO)*, Vicky Hogan (Health Canada), Leila Lackey (US FDA CBER), Shahrul Mt-Isa (MSD), Kitami Noriaki (PMDA)*, Shanthi Pal (WHO), Leo Plouffe (Bayer), Qun-Ying Yue (UMC)*, Cheryl Renz (AbbVie), Tomas Salmonson (former Chair CHMP), Stephanie Storre (Swissmedic), Sabine Straus (MEB, Chair of PRAC), Stephanie Tcherny-Lessenot (Sanofi)*, Steffen Thirstrup (CORS), Mariko Tsukuda (PMDA), Sebastian Vulcu (BI), Julie Williams (MHRA), Hong Yang (US FDA CBER), Xi Sherry Zhang (Gilead), Sanna Hill, Hervé Le Louët and Lembit Rägo (CIOMS), and Panos Tsintis (CIOMS Senior Adviser).

Apologies: Sergei Glagolev (Roszdravnadzor) and George Quartey (Roche)*.

Alternates did not attend: Karen Kaplan (MSD), Sara Khosrovani (MEB), and Hussein Laljee (Gilead).

* New to the working group since 1st in-person meeting.

NB: The Covid-19 pandemic put pressure on many WG members' time availabilities and several had to attend urgent meetings in parallel. For this reason, the order of reporting does not at times reflect faithfully the way discussions happened in practice.

Prior to the meeting

- The minutes of the 1st WG meeting held in Geneva, Switzerland during 17-18 September 2019 were approved by absence of comments.

DAY 1

Welcome and opening of the meeting

- Hervé le Louët, CIOMS President, welcomed the members, with special thanks to everyone for sparing their time during the Covid-19 pandemic, acknowledging that lives and jobs have been affected. Despite the crisis, CIOMS is continuing to work and Hervé thanked all for their commitment. The CIOMS WG XII is progressing very well, and we now find ourselves working on current-day, real-world issues. He suggested that we take this opportunity to work, not only on

innovative drugs, biologics and others, but also on the pandemic situation. He urged all to assess properly the benefit-risk (B-R) of drugs when taking into consideration the long-term nature of science, the mid-term nature of medicines, and the short-term nature of communication. This is a complicated issue of utmost importance. He then excused himself to attend an urgent crisis unit meeting in Paris, urging all to stay safe; and finished by saying he hoped to see everyone soon.

- Lembit added his words of welcome, commenting on the unusual circumstances, and how we are all learning a new way of working remotely. He thanked everyone for devoting their time during the difficult Covid-19 pandemic.
- The agenda and the working mode were adopted.
- Sanna requested consent to record the meeting for the benefit of taking minutes; passed on apologies from WG members unable to attend; and introduced new members.
- Vicky added her words of welcome and commended all on their work to date.

As the meeting took place during the ongoing Covid-19 pandemic, with several vaccine studies and potential treatment trials in progress – some being conducted better than others – the discussions resulted in a plan to issue a rapid release statement from the CIOMS WGs. The objective will be to promote a robust B-R approach, i.e. a scientific approach, in times of emergency to support the regulators and decision making agencies, despite the pressing difficulties faced during a pandemic. The statement discussions permeated the meeting but have been grouped at the end of this document.

[Subgroup 3 presentation: Benefit-risk landscape, chaired by Vicky, presented by Leila](#)

The Subgroup 3 PowerPoint slides and an in-depth outline document are available at the CIOMS website for WG XII, under password-protection.

Discussion

- In the light of what Hervé said, Lembit suggested addressing the current Covid-19 situation in the introduction, perhaps under “how to carry out B-R assessment during emergencies”. Emergencies are starting to be an important part of our lives, considering we have recently experienced SARS, MERS, Ebola, and now Covid-19, and more may still follow. How can we prepare for the next emergencies?
- Panos mentioned a recent FDA communication on additional risk minimisation considerations whereby physicians and patients are to make B-R decisions on whether to go to a hospital, and accept potential exposure to Covid-19, or to delay testing, e.g. laboratory tests.
- Leo added that the pandemic provides tremendous opportunities for learning and mentioned two elements:
 1. What happens when we do not have the usual safeguards in place in development: some are proposing using medications outside standard dosage and/or in combinations that we would normally not approve;
 2. The attitude that you may as well take a chance because the virus will kill you anyway. Some examples could go in an appendix.
- Richard feels this will be a difficult topic to address in terms of B-R assessment, and especially quantitative B-R assessment, because e.g. we have limited, poor quality data, and it is difficult to predict the course of the infectious disease; and therefore, inherently, there is going to be a lot of uncertainty. There is value in addressing this topic but it faces a lot of challenges.

- Lembit suggested we point out why things do not work and suggest what can be done to make B-R assessment in emergencies easier and more effective. Currently, there are a huge number of clinical trials underway for repurposing drugs, and although many are registered, they do not all go ahead, and those that do go ahead, often end up with relatively few patients and use different study designs, such that there is a great variety of trials taking place and pulling together conclusions is difficult. Maybe the major regulators, industry participants, and those experienced with running clinical trials, could do something preventatively to create standardised designs and/or a toolbox for emergencies, even though this is strictly speaking outside the mandate of this WG. This could result in better quality data i.e. indirect things that can facilitate B-R assessments during an emergency. What other things could we do differently to get a better starting point for B-R assessment in case of emergencies?
- Stephen advises including an example of R-B in the appendix but not too much in the main report body text.
- Sabine agreed it would be good for the WG XII to address Covid-19 in some way, preferably at a high level, e.g. in terms of lesson to be learned, which might be helpful for the next pandemic. On the one hand, the WG XII report will need to be sustainable, and not much about Covid-19, but on the other hand, when drafting a document like this during a pandemic, it would be fitting to address it in some way. It might be suitable to include something in the introduction, and then consider if we want to develop it further. Maybe the hydroxychloroquine data could be used for a lessons learned section?
- Tomas reminded the WG that decision making under time pressure was already discussed at the first WG meeting. The underlying issues include: transparency, conflicting messages, trust in regulatory processes, and semi-experts involved with B-R decision making.
- Richard mentioned there exist different philosophical approaches to decision making:
 - The FDA approach tends to maximise the effect of value - a utilitarian approach;
 - Rights-based issues - options available in line with the rights of the stakeholders;
 - Minimisation of potential catastrophic loss – when we predict a million or more deaths.
 The option offering the best expected value may not be chosen owing to the other considerations. Richard suggested including such philosophies for the report and can provide some book references.
- Patrick agrees that a different circumstance requires a different decision making philosophical model.
- Hong said subgroup 2 has also addressed emergency decision making to some extent and she sees two potential areas in the report for adding more information on the subject:
 1. The B-R framework - which part needs to be enhanced and/or can handle more flexibility?
 2. The methods section - what kind of clinical trial methodology design approach is possible; and what kind of tool can be used to communicate the approach?
- Vicky suggested to let thoughts mature on how to include emergency situations in the introduction e.g. whether it be integrated under transparency and trust, ways to deal with uncertainties, or lack of data. Readers would expect the report to address emergencies in some manner and we could provide some high level recommendations. We are still learning about Covid-19 but we can use our knowledge and rationale, and mention e.g. adapted clinical trial designs.
- Panos suggested placing risk management in the framework, perhaps before licencing, probably under the “Shift from post-market to lifecycle approach”. Similarly, Subgroup 2 would also need to include it under their topics. This is a relatively new process but there is not necessarily a new method yet.

- Leila reminded all to take care over the different regulatory contexts as they write.
- She also asked if there are other changes from the CIOMS IV report to mention? She has noted: risk management, emergency situations; change from covering only the “post-market” phase to covering the full “life-cycle”; patient level integration; adding a purpose statement that covers the main thrust of the document.
- Leo suggested adding:
 1. The evolving role of real-world evidence (RWE), as this can add to both sides of the equation: benefit and risk; and also provide insights into populations. There is rapid evolution in this area e.g. with regard to electronic health records and the contribution of RWE to the assessment of efficacy and safety. Mariko suggested that maybe RWE in decision making can be mentioned only briefly because this is going to be covered thoroughly by CIOMS WG XIII. Lembit added that the CIOMS WG XIII has just started, and that at a later time, we will perhaps make a connection between the two groups.
 2. The types of uncertainty e.g. statistical uncertainty about what we gain in the controlled studies, but also e.g. long-term uncertainty related to cell and gene therapies - the unknown unknowns and how this impacts B-R. In terms of risk minimisation, we know if we all commit to following patients for 10 years or more through a robust methodology, it will help us to elicit the unknown unknowns. Richard added that in the context of different types of uncertainties, sometimes that is dealt with by trying to model the worst case scenario, thereby choosing the 95th VaR as the worst outcome, at each link in the chain, which can lead to excessively conservative risk assessments.
 3. The evolving role of artificial intelligence (AI) and assisted decision making. There may be an opportunity to highlight that this is an evolving area. Richard mentioned a case example involving AI in a device: the device was used for information gathering, and contained an embedded algorithm, accrued by a regulatory agency, which was still evolving when in use in the real world by gathering more data. The B-R assessment in such a situation seems particularly intriguing. This is device-centric and perhaps too specific?

[Subgroup 1 presentation: Methods, chaired by Scott, presented by Richard](#)

The Subgroup 1 PowerPoint slides are available at the CIOMS website for WG XII, under password-protection.

Discussion

- Stephen feels there is a need to take particular care with vaccines or medicines given to healthy people for prevention, as the B-R considerations are often subtly different and not necessarily always only derived from within the trials. The harms become of greater importance when vaccines or medicines are given to healthy people than when they are given to people who are not well. So, the B-R balance needs to be adjusted for that, not necessarily in the calculations but in the thinking, and this is pertinent in the Covid-19 environment.
- Richard agrees. This is addressed at the biologics side at the FDA, including giving vaccines to healthy children to reduce the risk of adverse outcomes that might come later, and there is very low tolerance for any safety concerns. It is different with e.g. CAR T-cell therapies for cancer patients, when no other treatments have been successful and the patient has only weeks left to live. Then there is a higher risk tolerance. Hong agreed and added this will be included in the structured B-R framework too.

- Richard said subgroup 1 will cite Kahneman & Tversky prospect theory for the theoretical foundation.
- Panos asked Stephen about putting the number needed to treat (NNT) and the number needed to harm (NNH) in context. Is there a possibility of combining them in a measure? Stephen did not recommend using NNT under any circumstance, feeling it is used as if it were a pure number whereas it is not. Some speak of NNTs for treatments over one year and others over five years. When dealing with statins, advocates may talk in NNTs over five years, and even over 10 years. Just using NNT has no meaning. We need to have the follow-up time, and it is better to use absolute rates. If everyone used NNT over one year, it would be acceptable. In terms of combining them, Stephen did not think we can do that; it is better to look at the absolute risks; but if really necessary, they can be combined into a single measure using the Win Ratio.
- Scott said sometimes the separate marginal analysis of each outcome may not tell us the full story. Just analysing efficacy in one bin and safety in another bin, can have us miss important signals. It is important to consider them in-patient.
- Scott went on to elaborate, the reason we measure outcomes in trials, is to know how the patient is doing, but the way we analyse the outcomes, is upside down. We use the patients in the trials to analyse the outcomes, and by doing that, we can lose important information. We need to get the order of operations right. Outcomes need to be measured as they are experienced by the patients, they may be correlated or not correlated, positively or negatively; often we do not look. To compare two treatments, we need to look at the outcomes within-patient, then aggregate what is happening over the two different treatments respectively, and then make a comparison. That may be a better reflection of how two treatments compare.
- Both subgroups 2 and 3 are interested in including a heat map in their chapters to help visualise over time when an individual patient experiences success, or a risk, or both.
- Stephen mentioned the importance of discussing timing – if we were to do a metro analysis of surgery and base our outcome at the end of three to seven days, then all surgery would stop, because there are short-term risks, whereas the benefits are long-term. Usually, we do not get such a dramatic change over time for medicines. Particularly with preventative things, there are short-term harms and long-term benefits. If we balance the benefit and the harm at the wrong time, we will make major errors.
- Panos went on to say that in a trial, we can expect some of the safety aspects to come in the short-term and the B-R balance may change over different time points. He agrees it would be helpful to have some visual representation of how it could look for individual patients. It is novel to assess B-R within-patient at different time points. It would be good to choose some endpoints that combine at least two different B-R balances over time.
- Scott added that in trials, we choose efficacy objectives and outcomes, as well as safety outcomes; but we would like all to think about B-R as beyond a statistical analysis issue, and rather as a design and conduct issue. If we start to evaluate B-R throughout the product life-cycle, we should also think in terms of a study life-cycle. We can think about analysing the patient, and defining the efficacy and safety outcomes, but maybe we need to think about defining B-R outcomes at a patient level.
- Richard mentioned that one of the problems often faced in B-R assessment is that the data has not been collected in such a way as to make possible the patient-level assessment. This needs to be pushed back to the very early stages when studies are being designed, so that the data is available later. Sometimes, we may want to look at possible correlations between the benefits and risks at the patient level, but if we do not have data, we may be left with only the marginals.

- Hong highlighted that we need to think creatively about how to update the B-R bridge over the pre-market and the post-market phases. With many clinical trial designs, we have a limited time duration, and the B-R outcome may not be observed in the duration, especially when the risk is a long-term risk or the benefit is a long-term benefit. We cannot extend clinical trials forever and so it is important to update findings after approval.
- As mentioned, the Win Ratio can be one approach to prioritising outcomes, but Scott has also been working towards establishing an overall ordinal outcome that would encompass various elements e.g. in cardiovascular event prevention trials. We can have up to five levels of an ordinal outcome – we call this the desirability of the outcome – where in the most desirable category, the patient is alive and does not experience any events (strokes, myocardial infarctions (MIs), and bleeds from toxicity), and where in the least desirable category, the patient dies. The layers in between are defined by the number of events the patients experience and the consequences of those events. We can look at the distribution of the desirability of outcome ranking, an ordinal scale that would encompass that information, and see how treatments compare with that distribution, and that can compose the efficacy aspects of strokes, MIs, major bleeding, and mortality. This idea needs to be tailored to a disease area and how we compose the information to be measured (this is the biggest challenge). On the one hand, we seem to know how to analyse thousands of patients, but when we get down to analysing one, we seem to struggle. If we got better at this, we could have a better idea of the effects on patients.
- Panos asked whether the methods apply only when we look at hard outcomes, e.g. not all hypertension trials look at MIs and strokes. Can we apply the same method to different types of trials? What about the fact that many of the trials and programmes seem to consider rather the surrogates? Scott confirmed the concepts still apply. What information we compose and how we compose it will need a lot of work, and has to be tailored to the objectives and goals of the study and development. The framework can be used theoretically for any outcomes, even if they are surrogates or other. We have applied this to e.g. looking at dose-response patterns. When we are trying to figure out a dose, in many ways it is a B-R question: if the patient does not get enough, maybe they do not get the benefit, but if they get too much, it is going to be toxic. There may be pharmacokinetics (PK) parameters, or other parameters, that measure outcomes.
- Scott heads up an NIH (National Institutes of Health)-funded network on infectious disease studies, particularly in antibiotic-resistance areas, and in this setting, the team is trying to come up with a standardised B-R outcome in each of the major types of infection studies, based on the site of infection – Hospital-Acquired and Ventilator-Associated Pneumonia (HAP/VAP), blood stream infection, injury/abdominal infection, skin infection – and how we might construct it and standardise it across studies. There is still a lot of work to do on how to compose information, but theoretically, the construct and the idea (i.e. the concept of composing information), can apply to any outcomes.
- The subgroup 1 team would like to hear how much the WG would like them to cover the basic methods, e.g. effects tables, some of the more common quantitative approaches and qualitative approaches.
- Cheryl confirmed this chapter will be particularly useful in highlighting that there exist a variety of methods. In areas of quantitative methods, it would be helpful to identify the very best approach, and when it should be used, e.g. multiple-criteria decision analysis (MCDA), and some of the more commonly or increasingly used quantitative methodologies in the B-R space, when best they can be used, how to interpret the results, and so forth.
- Cheryl asked whether subgroup 1 will explore some of the newer applications of the patient preference studies and the different types of approaches to patient preference evaluations that

are now being applied in the B-R space? At the first WG meeting, we discussed inviting someone to participate in one of the subgroups on this topic. Richard agreed it is important to highlight the value of the patient perspective, but he is not sure how far this chapter should go into the methods for eliciting patient preference information. He is open to including it in this chapter or elsewhere in the report. Leo agreed it is an important element to cover. Maybe we can see how well we could coordinate with output from IMI Prefer, which focuses on patient preference studies.

- Stephen commented that we will need some good instructive examples to show both what can be done and the difficulties of ranking among the many, possible harms. There may need to be some ranking of severity, and doing that in a way that captures the variety of patient and clinician opinions will not be easy. Some problems are extremely difficult to solve in practice.
- Shahrul pointed out that we have not touched very much on the subgroup analysis of B-R balance within subgroups. The guideline from the EMA on subgroup analysis asks that the B-R should be specific to subgroups, so there should be methodologically more to it rather than just subset in the population. Should this be considered here? Scott feels subgroup evaluation should be based on B-R evaluation, but it is often based on either a subgroup that wishes to benefit from an efficacy perspective, or a subgroup that wishes to avoid a particular safety problem, but neither one is the group we want to treat. We want to treat the group that has both of those conditions ideally. Shahrul suggested expanding on this, especially in terms of the labels and reimbursements aspects. How do they come into play when we are only restricting to certain subsets of the trial populations?
- Stephen feels there is a danger in looking at subgroups and just marginal values, as this way, we can be subject to the Simpson's Paradox. We can end up with misleading results unless we look at the individual patients in the subgroups, defined by clinical characteristics, e.g. the elderly or the people with ST elevation, rather than at the subgroups which appear to have benefit, or harm or no harm. Those subgroups are extraordinarily dangerous. But if we define them by clinical characteristics, and if we just look at the marginal totals, it can be misleading too.
- Lembit thanked all for the great discussions.

DAY 2

Subgroup 2 presentation: Structured descriptive assessment, chaired by Patrick, presented by Cheryl

The Subgroup 2 PowerPoint slides are available at the CIOMS website for WG XII, under password-protection.

Discussion

- Patrick commended subgroup 2's advancement. Structured B-R is a very important tool and promoted by US FDA from early stages. At Pfizer, it is used extensively to prepare for regulatory submissions, but also for life-cycle management and portfolio prioritisation. Its use at all stages of product development is very important, even at pre-clinical development, and it is connected with guidance for risk management. There is a lot of linkage, a lot of opportunity to utilise the tool, being a practical tool for all in drug development and regulatory approval.
- Cheryl explained she has removed the word "descriptive" from "structured descriptive assessment", so that the "structured assessment" can be used more broadly, especially under the current circumstances.

- Cheryl felt that if done correctly, a structured approach should be applicable in any situation: vaccines, urgent situations, biologics, etc. This is an area where case examples can help to highlight how a structured approach can transcend any circumstance if used correctly, and by using it, we can increase the transparency and consistency of the B-R assessment.
- Cheryl asked if the WG recommends including other topics under subgroup 2?
- Panos commented that the CIOMS WG IV report has an outline of how to put together a B-R assessment – a list of key headings and some comments – and he recommends doing the same in the CIOMS WG XII report. The ICH clinical overview structure does not include risk management, but the FDA’s version and others do. It would be important to do this in a wider sense.
- Cheryl would like to create a simple tool, like a template, and then walk the reader through the various components, explaining what the purpose is and what the key information are. She would like to have the chapter used as a backbone for structured discussion, decision making, documentation and communication between stakeholders, e.g. regulators and industry.
- The subgroup 2 version is not constructed exactly as the ICH version because the subgroup 2 team believes that the simplified components should be the essence of the structured assessment, and that can make this chapter stand out from other guidances. The team would like it to be used throughout the life-cycle, not just for one document, e.g. a clinical overview, but different documents throughout the life-cycle.
- Stephen has a slight concern with putting absolute versus relative risks and benefits, and risk measures composites, a little further down, because they are quite important with relation to subpopulations. Stephen wonders if we want to be able to say something about the general principles about how we evaluate benefits and risks in subpopulations, although this may not be the appropriate place in the guidance. Maybe in an earlier chapter? If we have constant relative risks, then different subgroups, which are at different baseline risks, will have different absolute risks, and that is a very important point about any subpopulation. Cheryl said some of the key principles, e.g. the absolute/relative risk point, can be mentioned in the structured framework, even if it is expanded elsewhere. We can link the various places.
- Cheryl mentioned that key risks are defined here for the purposes of B-R assessment, and that this may differ from the identification of risks for the purposes of a risk management plan. E.g. something may be identified as a key risk for a product based on a patient’s perspective but we may not have the same as an important identified or an important potential risk in the construct of a Risk Management Tool (RMT) for the purposes of a risk management strategy or a risk management activity. When we define a key risk, we will need to emphasise this.
- Leo and Scott both agreed that the subgroup 2 approach is very good but both question whether there are ways to supplement it with some of the concepts on how risks and benefits may be correlated or associated, and how they might accumulate within a particular individual, in order to have continuity with the other sections. Cheryl said subgroup 2 would like to introduce the components of the structured framework, put some points to consider under each one, and guide the reader to think about key benefits and key risks on a population basis, an individual basis, etc, and link to other places in the report for further information. The team would like to introduce different ways of thinking about benefits and risks: as benefits, as risks, together, in patient groups, in individuals, etc. Scott wonders if we look at this major section, whether it is a product profile, based on indication, and components and some other things, and there might even be a couple of subsections:
 1. Devoted to some of the traditional summaries around typical approaches where we summarise the benefits and summarise the harms, one outcome at a time;

2. Clearly shows the thinking about it the other way, combining benefits and risks within patients, i) using patient data to focus on outcomes, ii) using the outcomes to figure out what is happening to the patients.
- Cheryl pointed out that even conveying the concept of what is a benefit or a key risk is complicated, and product development teams, who work on a product for multiple years, can struggle to say what they believe are the key benefits and risks, and how to combined them for the accumulative information within a population or an individual patient, especially with the ever-increasing amounts of data.
 - Leo suggested including in the subgroup 2 section the following:
 1. Types of uncertainties for informing decision making;
 2. Patient input - the nature of patient insights and how they inform decisions.

Cheryl said, at the moment, patient input is included under “special topics”. Leo feels by the time the CIOMS WG XII report will be published, patient preference will need to be included. Cheryl and Leo discussed including it under the product profile as a source of key evidence or under current treatment options. Hong made the point that patient input can be included under different components of this framework, e.g. different patients may have different opinions about the condition analysis, treatment options, burden of treatment, and the product profile. This is why patient input is probably difficult to put under one subtitle and why we include it under a separate chapter. Cheryl will think about how patient input can be mentioned in the construct of the framework component, and Hong’s point about how we can apply it across the life-cycle, the applications of the framework.
 - Leila suggested it might be useful to talk about the diversity of decision making models, as Richard mentioned, i.e. the decision making model in the context of utility maximisation versus rule-based. This may be best placed under the “responsible party”. We also need to mention about whether we are looking at a single decision maker, a group consensus, or just a group providing recommendations, etc. In addition, under “expert consultations”, if you are going through the effort of consulting experts, maybe add that you can also consult patients. Consulting a single patient does not represent all patients, it is not a substitute for a rigorous collection of patient experience and preference data, but it could be a useful addition to the process.
 - Patrick said with reference to 4.2, i.e. “cross-functional benefit-risk assessment team”, this role could be carried out by the safety management team (SMT), as introduced in CIOMS WG VI report – a multi-disciplinary team in charge of developing and analysing the structured B-R assessment. We should avoid multiplying the number of bodies dealing with a product because there is not only a clear linkage between the different elements that contribute to B-R, but also we need to have one team taking ownership of the B-R profile and working with the product along the life-cycle. It would be useful to provide a definition of SMT, even if it needs to be expanded with new function.
 - Cheryl felt it would be useful to discuss how to go about team input decision making, governance etc, and how to define that separately from safety teams. We would also need to discuss the remit of what is different between a B-R team and a safety team.
 - Patrick said one key output of a SMT is a B-R evaluation. At Pfizer, this is done e.g. with the Development Safety Update Report (DSUR). Cheryl will look at how to interface and handle this expanding remit for the structured safety teams. It seems different companies do this differently.
 - Cheryl mentioned it might be best to provide options and solutions, and then one of the solutions can link back to the SMT, and for companies that may be organised differently, we can

outline the role: the expertise needed on the cross-functional teams. We can at least highlight this construct of the various key functions, and companies can go about this differently.

- Vicky suggested that under the “special topics” section, it might be a good place to talk about emergency situations; and under “sources of data”, where in the early overview there was a long list of data sources, it may be good to divide the list into high quality evidence, medium quality evidence, and low quality evidence.
- Leila asked if we should suggest when in the life-cycle, and for what products, the process should be applied? Is this done for every product in the design phase of pivotal or confirmatory trials? Do we want to make recommendations? The first time it is done for a product will be the hardest. When should this be? What do we want to say about when it should be revisited?
- Cheryl went on to say there will be different applications for the structured approach at different times in the life-cycle. It would be a lot for a company/regulator to do at each milestone for every product. We might highlight what could be accomplished at each point in the life-cycle using the structured approach. What information is most important to inform decision making at a pre-clinical versus a registration timing? It will be different but we could highlight what will be the utility of the structured approach pre-clinical v another phase. Whether it will need to be done at each milestone is probably product-dependent. At least we can point out what we could get at each point in the life-cycle.
- Eventually, evidence generated at the different points along the life-cycle will go into the B-R assessment. We will not do a complete B-R assessment at each stage i.e. at the beginning because it is impossible to have enough evidence available.
- Patrick said that at Pfizer, the B-R governance process participates in the decisions to move products along the development cycle. Usually, they start doing the structured B-R assessments as soon as a product moves to early development. It is done first in an informal way, but to move further, they require a structured B-R assessment. Another critical part is vis-à-vis the competition (class and alternative available treatments for a medical indication, with the same mechanism or a different mechanism), and this is also part of the assessment. The SMT is put in place as soon as a product moves to testing involving humans.
- Cheryl would like to highlight where the visualisation would be best used in the structured framework but leave the details about the merit, the different types, and the approaches to subgroup 1. It may be possible to use some of the same examples in the introductory chapter, then use the same example in subgroup 2 for an application, e.g. if the introduction describes the purpose of the tornado plot under visualisations in subgroup 3, then subgroup 2 could show in the structured descriptive framework where it can be useful.
- Leo agrees and made points:
 1. At Bayer, B-R assessment is done prior to first-in-man trials, and the teams are encouraged to consider doing an update at each step of the way. One positive side-effect has been to inform other groups, including the marketing colleagues, about the B-R framework and helping to drive decision making based on the framework. It has been a very positive evolution and may be in itself worth considering.
 2. Visualization is only one aspect of communicating. Leo wonders if the subgroup 2 team has considered having a specific element about communicating B-R to different stakeholders. Probably, we cannot communicate in the same way to a regulatory agency, the clinical community, and a patient group.
- Leila added that for an organization that does not think about this carefully, it may appear to be a hard thing to do, and there will be a learning curve for organisations. It might be useful to provide a high level example of what Pfizer does, to show it is possible to start early and do it

throughout the life-cycle, and update it along the way. But at the same time, we may want to stress that even if you do this only once in the entire life-cycle of your product, it could still be useful. Leila added that if a company is only going to do it once, they should be told the junctures where it could be particularly useful, e.g when designing a pivotal trial or in a post-market setting when you have a new safety signal for a very serious safety issue.

- Cheryl concluded that this will be a practical way to present this to an audience of different corporate backgrounds.

General discussion

i) CIOMS XII guidance to supersede CIOMS IV guidance.

Lembit confirmed the collective vision seems to favour the new report superseding the CIOMS IV report. We will add a note on the CIOMS IV electronic copy to say it will be superseded, and where it is made available on the CIOMS website, we will add a disclaimer to the same effect. The WG agreed to this approach.

ii) Types of examples to be included in the guidance and the specific case studies to be used for each example.

iii) Develop a 'standardized' template or approach for examples so there is consistency – whose action?

Vicky would like to agree a shared structure / template to achieve a common look and feel for case studies e.g. giving key takeaways. The WG felt that examples can be quite diverse and we can work at the editorial stage to uniformalise them. Longer examples can be put in the appendix. The WG agreed to return to this at a later date.

iv) Glossary to be added to the guidance – whose action?

The CIOMS WG XI glossary team is in the process of preparing two glossaries:

1. The CIOMS Cumulative Pharmacovigilance Glossary (currently in draft format)
This includes all the pharmacovigilance terms used in the CIOMS reports: some have been adopted from other sources; some adapted from other sources; and some created;
2. The CIOMS WG XI Report Glossary (plan language) (currently in draft format).

- Sanna to share both drafts with WG XII, after some edits have been completed.
- At a later date, the WG XII can discuss if its report will need a glossary, and probably it will, we can then adopt what is fitting from the work done by WG XI and build on that.
- Vicky welcomes taking on board the WG XI glossaries, and she suggests as we write, we add new terms as we go along.

v) How to deal with references - reference by chapter / section or at the end of the document?

- There is no single CIOMS style. We can have cumulative references at the end of the report, which enables removing all duplication, or we can have references at the end of each chapter. For the time being, to enable us to move sections of text around as freely as possible, Sanna would like to ask all to put references in footnotes.
- Richard asked if there is a preferred reference style to use and reference management software e.g. EndNote, or other open source options?
- Shahrul suggested asking within the WG what reference management software they have, and then use that, and it may be easier to consolidate at the end.

Any other business

Editorial group

Usually, towards the end of a CIOMS WG's drafting process, we form an editorial group, with technical input and knowledge, including representation from each of the subgroups and a balance across regulators, industry and academia. The editorial group works with support from CIOMS secretariat (Sanna), and each WG subject matter tends to require a unique approach. Typically, the process will have Sanna compile the content, the editorial group will share out various tasks, or work in parallel; the group will feed back to Sanna or to the writer in question, we perhaps repeat this process a few times, and then lastly, the combined content is distributed to the whole group for approval.

For the coming 12-18 months, the WG is welcome to work fairly organically, each chapter team in their own style, such that each team can go through their thinking process in a way that works for them; allowing for how each person works individually. Different teams tend to work at different speeds and with different group dynamics. At a later stage, we can look to uniformise the output.

Covid-19 situation

Richard mentioned that the Covid-19 situation has some serious impact on the bandwidth at the FDA, and although the FDA is committed to this WG, there may be some unexpected demands on time and priorities that will need to be balanced. Lembit appreciates some are heavily involved with Covid-19, understands, allows for flexibility, and this information is well taken. Vicky agrees.

CEPI (Coalition for Epidemic Preparedness Innovations)

Stephen related that he is the statistician to a meta-data safety monitoring board for CEPI; which had been preparing to coordinate a series of trials of different vaccines in a potential future pandemic before the current pandemic took effect. CEPI is not tied to any one company, government, or country, and has set up a very good structure to try to coordinate these trials of vaccines. It seems there could be a role for CIOMS, or some other body, because such coordination does not seem to be happening with medicines in the same way.

Next steps

Vicky would like to set up a progress update conference call towards mid/end of July in order to prepare for the next face-to-face meeting in September. She will keep an eye on the evolving pandemic and will remain open to feedback from the Co-Leads as to their availability to participate.

The next CIOMS XII working group meeting is provisionally scheduled for 8-9 September in New Jersey, USA, hosted by Pfizer.

Action items and conclusions

- All subgroups will need to continue working closely together to link together subjects and to avoid overlaps, e.g. in the areas of life-cycle and data sources.
- Vicky confirmed she is impressed with the amount and quality of work done, and she is pleased to see how the different subgroups are starting to coalesce with their thought processes. She did not see many areas of overlap between subgroups. There has been lots of good thinking about the individual topics and which subgroup is primarily responsible for covering a given topic, and it shows the early planning has paid off. The two days have been very productive, and it is great to see how far we have come in just over seven months. Thanks for everyone's efforts despite the Covid-19 crisis.

- Lembit is also very grateful for all the efforts. This is a very productive group and it is enjoyable to work together. He looks forward to a valuable result at the end.

CIOMS WG XII Rapid release statement

Background

- In the current the Covid-19 crisis, even the best academics are challenging the accepted methodologies for B-R evaluation. There is a role for CIOMS to re-establish the role of methodology in the way B-R analysis is conducted, even in a crisis. The same methodologies that are used in pre-marketing and post-marketing to support regulators and decision making agencies are still relevant during pandemics. A structured/rational approach must be applied, not an emotional one. It is important to promote transparency in times of emergency where there is a particular flow of information and misinformation.
- WHO issued a press release on 29 April, summarising available treatments, and it provides some coordination. Recently, the *New England Journal of Medicine* published an article about being careful to not lose reason, as everyone is interested in efficacy, not thinking about safety.
- In the cases of chloroquine and hydroxychloroquine in combination with erythromycin (Pfizer product), a global survey result shows they are the first drugs used globally to treat the early and late forms of Covid-19. Regarding the efficacy of these products, the case was made very quickly and adopted by chief scientists and politicians, bypassing the scientific channels, and people have lost all references on how we do an evaluation on drug efficacy and safety. When hydroxychloroquine started being promoted, studies were only concerned over efficacy: the first study in China was based on only 20 patients; subsequent studies were questionable; there was a study showing no effect; a study with 1,000 patients that had a questionable methodological standpoint; and people completely neglected the effect of QT prolongation. Even the efficacy criteria that are normally applied to viral infection studies were not really applied. Safety considerations were put on the backburner to the extent that the FDA had to issue a reminder about the risk of combining chloroquine and hydroxychloroquine in combination with erythromycin. At the same time, we saw a surge in e.g. arrhythmias due to the QT prolongation effect when the two drugs are combined. More people may die from the drug combination issues than from Covid-19, at least the minor forms seem in Marseilles, from which patients may have recovered with no treatment.
- Are there ways to accelerate scientific judgement? Should we adopt a new set of rules? It cannot be left to people who have no expertise.

Statement to be issued collectively by CIOMS WGs

- It was decided that the CIOMS WG XII would start drafting a rapid release statement to establish the need to have rigorous scientific methodologies applied to a product B-R assessment, even if the circumstances may trigger different types of approaches.
- The rapid release statement will be issued by CIOMS on behalf of the following WGs:
 - CIOMS WG XI: Patient Involvement in the development and safe use of medicines;
 - CIOMS WG XII: Benefit-risk balance for medicinal products – Update of CIOMS IV;
 - CIOMS WG XIII: Real World Data/Real World Evidence in regulatory decision making;
 - CIOMS WG on Drug Induced Liver Injury (DILI);
 - CIOMS WG on Clinical Research in Resource Limited Settings.
- All the WGs above are tackling some aspects of the current emergency and each guidance document will include some elements to address the emergency situation.

- Any statement issued needs to support other high level structured efforts e.g. by WHO and the National Institutes of Health (NIH) studies.
- The B-R assessment is dependent on the data we have. Or, we may have bad quality data or no data. There is some work underway at the moment to gather quickly good quality data that would facilitate future assessments. We can give recognition to these efforts.
- A disclaimer will be included to say that this statement represents the consolidated views and opinions from the participants of the CIOMS WGs but does not reflect any views of individuals or member organisations where these individuals are employed or with which they are affiliated (in the style of DIA disclaimers).
- The decision to issue the statement by CIOMS on behalf of the five WGs was partly to allow a degree of separation between the statement and individual stakeholders and individuals.

Writing process

- The drafting process will offer an opportunity for the stakeholder organisations to consult internally and offer their comments for consideration, or to opt out if preferred.
- Lembit, Cheryl, Steffen, Patrick, and Leo volunteered to draft the statement, to be joined by Richard, following approval from the FDA leadership; Vicky and her group at Health Canada have offered to review the statement; and all stakeholders will have an opportunity to comment.
- Lembit will circulate a preliminary draft on 1st of May.

Roles within the CIOMS XII Working Group

Chairwoman: Vicky

Co-Chair: Patrick

Co-Chair: Scott

Subgroups and Co-Leaders (Co-Leaders' names are underlined)

Subgroup 1

Methods

Including integrated B-R methodologies / patient level

Shahrul

Leo

Panos

Scott

Stephen

Richard

Patrick

Qun-Ying

Stéphanie T

Subgroup 2

SDA

Structured descriptive assessment

Sherry

Stewart

Cheryl

Stephanie S

Sebastian

Julie

Sergei

Hong

Kitami tbc

Mariko

Subgroup 3

Benefit-risk landscape

Guacira

Takahiro

Steffen

Tomas

Leila

Sabine

George

Shanthi

Stéphanie T

Vicky is currently not attached to a subgroup.