Methodology Working Party (MWP) Interested Parties Meeting

7 June 2024 (09:00-13:00) Virtual meeting only, via Teams

General comment on the event

- Discussions will be assessed for input in development of EMA MWP workplan for next three years.
- No official EMA minutes were taken.
- Focus on statistical topics, but MWP's remit is broader.
- The format, with industry raising many questions and then regulators replying to some, did not allow any further discussion between the parties.

Agenda

Agenda MWP Stakeholder Interact

Session outline

Session	Talk	Topic	Company	Speaker
1	1	Causal inference in clinical trials	J&J	Sanne Roels
1	2	Covariate adjustment in randomised clinical trials	Novo Nordisk	Henrik Ravn
1	3	Multiplicity adjustments in multi-armed trials with inferentially independent hypotheses	J&J	Tobias Mielke
1	4	Designing clinical trials with a delayed treatment effect	Novartis	Dominic Magirr
1	5	Meta-analyses and the estimand framework	Novo Nordisk / AZ	Christian Pipper & David Wright
2	6	Augmented controls in the confirmatory setting	UCB	Marc Vandemeulebroecke
2	7	Synthetic control arm in clinical trials for regulatory and HTA/reimbursement decision making	Novartis	Björn Holzhauer
2	8	Dynamic borrowing for clinical trials	J&J	Roel Straetemans
2	9	Extrapolation across related diseases	UCB	David Wright
2	10	Dose recommendations across related indications	Novartis	Marina Savelieva

Session 1

1. Causal inference in clinical trials
Sanne Roels (Johnson & Johnson)

The causal inference topic was considered both timely and relevant, particularly for antidrug antibody ICEs. The estimand framework is method agnostic, leading to questions about whether specific guidance is needed. Similarly, the causal inference framework is not excluded by the ICH E9(R1) guidelines and can be applied when deemed appropriate. There was a reluctance to make general statements on the acceptability of the causal inference framework, specifically counterfactuals, on efficacy claims. However, opportunities were mentioned in the presence of ICEs. The focus was on more specific opportunities rather than general ones. It is important to decouple "estimand" matters from "estimation" matters. Throughout the discussion, it was felt that there might not have been a common, well-understood, and clear understanding of how a causal inference framework may resolve some additional issues on model reliance, particularly when a patient-specific prediction of the control group outcome was available (potentially built through ML techniques, so-called "supercovariates").

2. Covariate adjustment in randomised clinical trials
Henrik Ravn (Novo Nordisk)

During the MWP meeting, it was stated that covariate adjustment is of high priority and there were questions raised about the efficiency gains that modern methods can bring. However, there was also skepticism about data adaptive methods. Concerns were raised about how new methods affect the interpretation of results, especially if the overall population has different characteristics and whether synthetic covariates affect this differently. Decreasing the sample size if efficiency gains are expected remains a sponsor's risk. In general, using complicated methods to save very few patients is not encouraged. The preference is on well-understood methods rather than on less well-understood methods, although it was felt that additional clarity may help there too. There was a desire for understandable and explainable methods, but this may have limited use if the target is on a different causal contrast. The main interest should be in that causal contrast rather than the role of covariate X/Y in the same analysis, although this does not exclude the need to investigate the role of covariate X/Y, which should be targeted by a different analysis.

3. Multiplicity adjustments in multi-armed trials with inferentially independent hypotheses
Tobias Mielke (Johnson & Johnson)

MWP experts mentioned that the topic of multiplicity adjustments in multi-armed trials with inferentially independent hypotheses is of high priority and work is ongoing, but there is no established level for k-FWER control. The discussion started with a question on how to define inferential independence of research hypotheses. It was pointed out that even drugs with the same underlying mechanism of action may not make the respective hypotheses inferentially dependent, as those candidate drugs would alternatively be tested in separate trials. However, the discussion deviated from the original question and instead of discussing the need to control suitable error rates, considerations were brought up on how the k-FWER could be decreased by increasing control-group allocation. A point was raised that there is interest in controlling the posterior probability of making a type-1 error,

due to a concern that more less well studied drugs would make it to Phase 3 through the implementation of multi-armed trials. While there could be validity to this concern, it would rather point towards requirements on sufficient signs of efficacy prior to starting Phase 3, rather than raising a motivation for stricter type-1 error control in Phase 3. Lastly, the 2 trials rule was brought up as a reason why a larger k-FWER vs. separate trials might be appropriate. However, in this point of the discussion, it was not acknowledged that the 2-trial rules is not implemented across all indications and for all situations were multi-armed trials are of interest. In particular, multi-armed trials would be considered for situations, where it is already hard to recruit the required information for a single trial (e.g. due to costs or size of the population).

4. Designing clinical trials with a delayed treatment effect
Dominic Magirr (Novartis) & Carl-Fredrik Burman (AstraZeneca)

MWP expert stated that there is already a large simulation study on the overall topic of designing clinical trials with a delayed treatment effect (https://doi.org/10.48550/arXiv.2310.05622). The MWP has already decoupled testing and estimation in the benefit-risk setting and acknowledged the issue that certain types of weighted log-rank test statistics reward early harm. The MWP favored a pre-specified weighted test based on prior knowledge, as opposed to an adaptive choice of weights, and favored corrections for multiplicity when using multiple summary measures, stressing that they should be pre-specified. However, other MWP members were not satisfied with the decoupling of testing and estimation and were not convinced of the need to correct for multiplicity for multiple summary measures. Finally, another MWP member expressed concern that a weighted log-rank test could increase problems with informative censoring and was also skeptical about decoupling testing from estimation. E.g. in the extreme there could be a significant test with an ineffective effect estimate. On the other hand, he would not require a relevant effect along with a positive test.

5. Meta-analyses and the estimand framework
Christian Pipper (Novo Nordisk) & David Wright (AstraZeneca)

One MWP member noted that estimands do not need to be fully aligned, but different estimands generate heterogeneity. Another MWP member asked if this can be solved with more granularity in estimand definition, as they do not want to use this to be prohibitive for a meta analysis.

Session 2

6. Augmented controls in the confirmatory setting
Marc Vandemeulebroecke (UCB)

7. Synthetic control arm in clinical trials for regulatory and HTA/reimbursement decision making
Björn Holzhauer (Novartis)

6) and 7) were discussed jointly.

MWP expert informed that EMA is currently updating the Single Arm guidance. In the context of placing augmented controls in between RCTs and single-arm studies, it was highlighted that the single-arm study guidance puts emphasis on the choice of suitable endpoints. The pilot program on SA trials is still open and if it is not closed, it is not possible to have clear guidelines on standards

for controls. So currently, different standards are possible. After the pilot is closed, they will create guidelines. There was a call to submit projects to the pilot. One MWP member stated that the tension is between how to make use of data and not compromise the decision. The obvious benefits for him lie in answering other variables than the primary efficacy variable.

Another MWP member remarked that augmented data cannot replace RCT or concurrent control. However, there is an opportunity to use only concurrent control in the comparison, although it will be underpowered. The concept of augmented controls is more appealing than borrowing information from other trials. The difference between augmented and synthetic control, or SA, is that with augmented controls, there is a "self-standing part of the trial". Augmented controls can answer additional questions for registration and are considered better than borrowing treatment effect. However, they are not ready to set specific standards for requirements and there is a clear hierarchy in quality where augmenting with controls from RCT is better than from RWD. The MWP chair concluded that since the SA study was mentioned a couple of times, he felt the urge to mention that the SA paper starts from the Research Question and the governing principle is that the question needs to be asked "why does the research question need external data". He disagreed with the mentioned conflict between SA and ICHE10. For him, the SA paper is on the research Q while the ICHE10 is on assessment.

- Further comments:

  It was noted that the EMA is currently flexible on raw data access while their raw data pilot is ongoing, and more guidance may be developed in the future. Potential difficulties in physical ICF access are acknowledged. The discussion of the role of augmentation should consider not just the situation of primary proof of efficacy, but the full spectrum of possible applications, such as providing additional evidence of efficacy for the totality of evidence, benefit/risk, etc. It is important to keep in mind that regulatory decision-making is based on a broad evidence base. On the one hand, it's important to focus on the key question of interest, but on the other hand, it's also more than that. Augmented controls cannot replace RCTs when conducting an RCT is feasible. Yet, augmented controls are better than single-arm trials because you can at least look at some trial-internal comparative data while accepting a power loss in that subset. This may be acceptable for non-primary decision-making. Borrowing information on the treatment effect is more problematic than borrowing information on the control. With respect to using non-RCT-data, such as RWD or synthetic data, as a source of augmentation, there is an informal hierarchy of trustworthiness. It is important for sponsors to start from the research question and justify why it requires external information sources at all, rather than starting from the question of what data is available.

8. Dynamic borrowing for clinical trials
Roel Straetemans (Johnson & Johnson)

One MWP expert noted that he liked the concept of buying power in the part where you are interested in, at the expense of Type 1 error outside that range. On the question of exchangeability, he gave the somewhat strange opinion that it depends on the conclusion of the outcome. However, he still agreed that assessing the exchangeability assumption is a very important part and mentioned that they are reluctant to leave the decision taking to modelling. Another MWP expert stated that his reflection was that BDB is clearly not standard but there are cases where it is accepted. He mentioned some interdisciplinary discussion within EMA on BDB, specifically what

would the consequence be of accepting the approach. He also mentioned that a detailed discussion on OC is helpful for those internal discussions and made a plea to be clear on the OC.

Another MWP expert noted that the role of the study within the overall Clinical Development Plan (CDP) is crucial. When going for scientific advice, it is important to be clear on the framing of the study. BDB is on the workplan and a Bayesian guideline is being worked on (as well as ICH E20); these points will probably make their way in there. An EMA MWP expert concluded that it is hard to support assumptions and it is easiest when the augmented controls are from a comparable RCT. It is very difficult to prespecify criteria for exchangeability and hard to ensure that all choices were made prospective and not post hoc.

9. Extrapolation across related diseases
David Wright (AstraZeneca)

At the MWP meeting, an NCA expert mentioned that she felt the agency was clever to not mention explicitly pediatric extrapolation in the extrapolation paper, although she admitted this was in the back of her mind. For her, the source/target population boils down to the same as other disease/current disease. She believes that the extrapolation concept paper should hold and that acceptance for extrapolation will depend on the "impact" of the decision. She also mentioned the need to understand different sources of variability and to distinguish between the uncertainty that's there anyway, and the additional uncertainty brought in by extrapolation. Another NCA expert stated that for her, the main question is "how does the benefit/risk" look like. She also mentioned that it is crucial to make a distinction between symptomatic or disease-modifying drugs and that extrapolation is more acceptable for symptomatic drugs, in rare diseases with reliable BM across diseases. One MWP expert concluded that the extrapolation of safety profiles is different for clinicians.

10. Dose recommendations across related indications
Marina Savelieva (Novartis)

An NCA expert mentioned how she really appreciated the point of addressing uncertainty brought forward by the last speaker. She advocated for "bench marking" of what has been approved by what is being proposed. The current guidance in place is applicable not only to paediatric extrapolation but also more generally across indications. The NCA expert also mentioned that probably somewhat more transparent communication on why some applications were rejected would help drug developers to ameliorate the proposed extrapolation approaches. Another NCA expert noted that the points above about benefit/risk and symptomatic vs. disease-modifying were reiterated and that it is already happening now, but at sponsor risk.