Type I error Considerations in Master Protocols with Common Control in Oncology Trials: Report of an American Statistical Association Biopharmaceutical Section Open Forum Discussion[+]

Rajeshwari Sridhara[1], Olga Marchenko[2], Qi Jiang[3], Richard Pazdur[1], Martin Posch[4], Mary Redman[5], Yevgen Tymofyeyev[6], Xiaoyun (Nicole) Li[7], Marc Theoret[1], Yuan Li Shen[1], Thomas Gwise[1], Lorenzo Hess[8], Michael Coory[9], Andrew Raven[10], Naoto Kotani[11], Kit Roes[12], Filip Josephson*[12], Scott Berry[13], Richard Simon[14], Bruce Binkowitz[15]

[1] Oncology Center of Excellence US FDA, [2] Bayer, [3] Seagen, [4] Medical Statistics at the Medical University of Vienna, [5] Fred Hutch, [6] Janssen, [7] Merck & Co., Inc., Kenilworth, NJ, USA,[8] SMC, Switzerland, [9] Department of Health, Australia, [10] HC, Canada, [11] PMDA, Japan, [12] Swedish Medical Products Agency (MPA), [13] Berry Consultants, [14] Simon Consulting, [15] Shionogi

*The opinions stated here are those of the author and not necessarily of the Swedish MPA or the European Medicines Agency.

Abstract

This article provides a summary of discussions from the American Statistical Association (ASA) Biopharmaceutical (BIOP) Section Open Forum organized by the ASA BIOP Statistical Methods in Oncology Scientific Working Group in coordination with the US FDA Oncology Center of Excellence on October 8, 2020. Diverse stakeholders including experts from international regulatory agencies, academicians, and members from the pharmaceutical industry engaged in a debate on type I error considerations in master protocols with a common control. Although there were concerns in specific situations where type I error adjustment may be necessary, the

[+] This is an Accepted Manuscript version of the following article, accepted for publication in Statistics in Biopharmaceutical Research: Sridhara, Rajeshwari, et al. "Type I error considerations in master protocols with common control in oncology trials: report of an American statistical association biopharmaceutical section open forum discussion." Statistics in Biopharmaceutical Research 14.3 (2022): 349-352: https://doi.org/10.1080/19466315.2021.1906743. It is deposited under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike License (https://creativecommons.org/licenses/by-nc-sa/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

panelists agreed that adjustment of type I error for multiplicity when a common control is used may not be necessary if the hypotheses are inferentially independent.

Key words: oncology drug development, master protocols, common control, Type I error adjustment, inferentially independent hypotheses.

Introduction

The Biopharmaceutical Section (BIOP) of American Statistical Association (ASA) in coordination with the Oncology Center of Excellence (OCE), U.S. Food and Drug Administration (FDA) initiated open forum discussions on different aspects of statistical considerations in oncology clinical trials, aligning with the OCE's 'Project **S**ignifi**CanT**' (Statistics in Cancer Trials). These open forum discussions are designed to engage experts and diverse stakeholders who understand the unique aspects of oncology clinical trials. Issues discussed in these open forum meetings can inform design and analysis of future oncology clinical trials.

The first virtual open forum discussion on Type I error considerations in Master protocols with common control in oncology trials was held on October 8, 2020 (Sridhara et al. 2020). The panel consisted of diverse stakeholders and experts from international regulatory agencies, academicians, and members from the pharmaceutical industries engaged in oncology product development. The discussions were moderated by the co-chairs of the Statistical Methods in Oncology Scientific Working Group of the BIOP, Qi Jiang, Ph.D. (Seagen) and Olga Marchenko, Ph.D. (Bayer), and Rajeshwari Sridhara, Ph.D. (Contractor at OCE FDA).

The discussions focused on whether Type I error adjustments are necessary in a randomized oncology clinical trial under a Master protocol that is designed to evaluate multiple treatments versus a common control. In this report we summarize these presentations and discussions. The forum started with 6 formal presentations by covering introduction to the topic, and point-counterpoint views on Type I error adjustment from academicians, members from companies engaged in oncology drug development, and members from various international regulatory agencies. The 20 panelists for the discussion included members of the BIOP Statistical Methods in Oncology Scientific Working Group representing pharmaceutical companies, representatives from international regulatory Agencies (US FDA, European Medicinal Agency (EMA), Health Canada (HC), Pharmaceuticals and Medical Devices Agency (PMDA) from Japan, Australian Department of Health Therapeutic Goods Administration (TGA), and Swissmedic (SMC)), academicians and expert statistical consultants (see the agenda in Appendix).

Master Protocol

Master protocols are identified as protocols that try to answer multiple questions with respect to multiple diseases and/or multiple treatments (Woodcock and LaVange 2017; FDA 2018; Meyer et al. 2020). These protocols have been implemented as different types of trials such as basket trials, umbrella trials and platform trials depending on the Master protocol objectives. Master protocols have the potential to accelerate drug development and save resources, particularly patient resource, with centralized governance structure, data sharing and use of a common control in the development of innovative treatments for cancer patients. Generally, it is understood that in a Master protocol the use of a common control can increase efficiency. However, there are differing views regarding adjustment of Type I error for the multiple

comparisons of different treatments to the same control arm in a randomized controlled trial utilizing a Master protocol (Parker and Weir 2020; Bretz and Koenig 2020). In the next section we report the summary of the discussions.

Are Type I error Adjustments Necessary?

Academic Viewpoints by Panelists

Multi-arm trials with a common control can substantially increase the efficiency of drug development programs. In difficult experimental situations, for example, rare diseases this can enable investigators to address questions that may not be otherwise feasible. However, these trial designs have an impact on the statistical properties of hypothesis tests. For example, due to the shared control arm, the test statistics for the treatment control comparisons are positively correlated. As a consequence, the familywise error rate is smaller compared to tests in independent trials if an unadjusted 5% level of significance is applied. This, however, comes at the cost that the probability to simultaneously reject multiple null hypotheses increases compared to independent trials. Furthermore, current regulatory guidance requires study-wise error rate control. For multi-armed trials it has been argued that no familywise error rate control is necessary, because it would also not have been required if the trials were running independently. Still, for a sound interpretation of trial results as well as from a societal perspective, the quantification of the overall operating characteristics of a trial is important (Collignon et al. 2020).

When considering multiplicity adjustments, defining the 'family' of hypotheses for which to control the error rate is crucial (Howard et al. 2018). One option is to define criteria based on the type of study hypotheses. For example, adjustment of type I error for multiplicity when a

common control is used may not be necessary if the tested hypotheses are inferentially independent. Then control of Type I error for multiplicity at treatment or sub-study level may be considered sufficient.

To assess the overall operating characteristics, besides the familywise error rate one could consider approaches to control the false discovery rate (Wason et al. 2020), especially, when there are a large number of treatments, or, control of an expected loss (Collignon et al. 2020), or, use a Bayesian decision theoretic approach (Muller et al. 2007). In Master protocol trials, however, controlling the Type I error at the overarching study or platform level is methodologically challenging (Meyer et al. 2020; Posch and Koenig 2020), as the number of treatments or diseases to be evaluated may be unknown at the beginning of the study and corresponding flexible statistical methods are required.

Furthermore, there can be various other sources of multiplicity outside of treatment versus control comparisons such as multiple endpoints, interim analyses, adaptations, change of control arms, etc. which will require careful planning for multiplicity adjustments. Some of the biggest challenges in studies under a Master protocol are not only about the design of the trial or Type I error control, but also about the logistical and operational aspects of the study.

Industry Viewpoints by Panelists

Examining 'relatedness' or correlation of hypotheses and tests are important. Careful examination of multiple opportunities for a drug to establish efficacy should guide any adjustment for multiplicity (Stallard et al. 2019). Examples where Type I error adjustment would be necessary include comparison of different dose levels of the same drug to a common control, and factorial designs where comparisons of control vs. drug A, control vs. drug B, and control vs combination of drug A and drug B. Special considerations are needed when investigational

drug combinations are being evaluated as the relatedness introduced by some common drug component shared by investigational treatment regimens. While it may not be a Type I error issue, potentially an increase in the probability of multiple false positive results may be observed if the shared control under- performs by chance alone (Howard et al. 2018; Collignon et al. 2020). In general, no adjustment of Type I error is necessary if multiple experimental arms are distinct treatments and the decisions are independent (Bretz and Koenig 2020; Collignon et al. 2020; Howard et al. 2018; Parker and Weir 2020).

Often oncology studies are powered for the primary endpoint of progression-free survival and may lack the power to test hypothesis based on overall survival outcome within a sub-study. Typically, the null hypothesis of no treatment effect on overall survival is tested only after a successful rejection of the primary null hypothesis. Therefore, there is an additional safeguard for the control of a Type I error further diminishing the concern regarding the use of a shared control arm for the overall survival analysis. In such a case, there may be an additional rationale to pool data across multiple treatment sub-studies. However, the inferences will be dependent when pooling is considered.

Regulatory Agencies' Viewpoints by Panelists

When there is a clear clinical dependence, strong control of family-wise type I error probability is expected (Collignon et al. 2020). Interim analysis of a treatment versus control may potentially disclose information for another treatment to control comparison. In such a scenario multiplicity adjustment may be necessary. During the discussions not all regulators agreed, a few expressed the need for adjustment and opined that if control is shared then by definition the treatment to control comparisons are not evaluated independently, and when different drugs with the same mechanism of action are each compared to the shared control, the situation is similar to testing

different doses. It was also recognized the possibilities of errors made in clusters even if hypotheses are inferentially independent; however, this is not exactly the same as usual Type I errors and multiplicity correction may not be the solution in all cases.

In determining if Type I error adjustment is necessary when multiple treatments are compared to a common shared control, it is important to examine whether the experiment is set up to make independent inferences or not. When several treatments with similar mechanism of action are being evaluated with a common control, the key issue to be determined is if individual drugs are being evaluated independently or information is borrowed from one treatment arm to the other. If information is not borrowed, then adjustment for multiplicity may not be necessary.

Some panelists expressed concern about leakage of information regarding the performance of a control arm which could potentially affect future treatment arms that may be initiated.

It was acknowledged that current experience with confirmatory trials under Master protocol is limited. In the future with more knowledge and experience, delineation of when and how Type I error adjustment is necessary may become clear.

Highlights of Discussion

Oncology drug development is going through revolutionary changes both in terms of type of indications and type of drugs, and with these changes there are increased numbers of smaller molecularly defined subsets of patients or rare disease groups and, of unique indications, which pose challenges. Many drugs have been approved based on single arm trials and smaller number of patients based on tumor response. However, randomized studies are the key to ensure that the observed clinical benefit and risk are attributable to the treatment under consideration and randomized studies should be conducted whenever possible. Use of Master protocols with a

common control can allow the conduct of randomized studies in such situations. To date our regulatory experience in evaluating products based on randomized study under Master protocol for the treatment of cancer is limited.

A question regarding comparing one experimental treatment to another experimental treatment was clarified: in this type of Master protocols in oncology the contractual agreements in place do not allow for such comparisons and each treatment arm data belong to the respective manufacturer or sponsor. In addition, sponsors may have little interest in participating in clinical trials where their drug product is directly compared to a competing drug product of another sponsor.

It was recognized that it is not unusual for patient advocacy group to run a platform trial and be in control of the data. However, the companies or sponsors who have developed the investigational product can use the data and submit licensing application for marketing to the regulatory Agencies.

This discussion focused on the need for adjustment of Type I error when multiple treatments are compared to a common control in a study evaluating cancer drugs. While there were concerns in specific situations where type I error adjustment may be necessary, the panelists agreed that adjustment of type I error for multiplicity when a common control is used may not be necessary if the hypotheses are inferentially independent. However, when some of the hypotheses are inferentially dependent such as comparing different doses of the same drug or drug combinations with the same components, Type I error adjustment could be necessary. Clustering of errors is a consequence of the statistical dependence of the hypotheses tested. Its impact on decision making errors is not fully understood. Although clustering of errors might not be a critical issue,

an assessment of the overall operating characteristics can be an important factor for the interpretation of trial results, especially if a large number of treatments is studied.

It was recognized that in conducting a study under a Master protocol there are multiple, important logistical challenges beyond the considerations for type I error control. There are also multiple, important advantages – even if correction for type I error is required.

This discussion did not consider situations where there are staggered entry and exit of treatment arms and the use of non-concurrent control data. This aspect was planned to be the focus of discussion in a future open forum.

Acknowledgement: Authors thank Joan Todd (OCE FDA), Cindy Lu (Biogen) and Rong Liu (BMS) for taking the meeting minutes.

Martin Posch is a member of the EU Patient-centric clinical trial platform (EU-PEARL). EU-PEARL has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement Nº 853966. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA and Children's Tumor Foundation, Global Alliance for TB Drug Development non-profit organisation, Springworks Therapeutics Inc. This publication reflects the authors' views. Neither IMI nor the European Union, EFPIA, or any Associated Partners are responsible for any use that may be made of the information contained herein.

References

- Bretz, F., Koenig, F. (2020). "Commentary on Parker and Weir". *Clinical Trials*, 17(5), pp. 567-569. DOI:10.1177/1740774520941420

- Collignon, O., Gartner, C., Haidich, A. B., Hemmings, R. J., Hofner, B., Pétavy, F., Posch M, Rantell, K., Roes, K., Schiel, A. (2020). "Current statistical considerations and regulatory perspectives on the planning of confirmatory basket, umbrella and platform trials". *Clinical Pharmacology & Therapeutics*, 107(5), pp. 1059-1067. DOI: 10.1002/cpt.1804
- FDA Draft Guidance for Industry: Master Protocols: Efficient Clinical Trial Design Strategies To Expedite Development of Oncology Drugs and Biologics, October 2018, https://www.fda.gov/media/120721/download
- Howard, D. R., Brown, J. M., Todd, S., & Gregory, W. M. (2018). "Recommendations on multiple testing adjustment in multi-arm trials with a shared control group". *Statistical Methods in Medical Research*, 27(5), pp. 1513-1530. DOI: 10.1177/0962280216664759
- Meyer, E. L., Mesenbrink, P., Dunger-Baldauf, C., Fülle, H. J., Glimm, E., Li, Y., ... & König, F. (2020). "The evolution of master protocol clinical trial designs: A systematic literature review". *Clinical Therapeutics*, 42(7), pp. 1330-1360. DOI: 10.1016/j.clinthera.2020.05.010
- Parker, R. A., & Weir, C. J. (2020). "Non-adjustment for multiple testing in multi-arm trials of distinct treatments: Rationale and justification". *Clinical Trials*, 17(5), pp. 562-566. DOI: 10.1177/1740774520941419
- Posch, M., & Koenig, F. (2021). "A Comment on "The Role of p-values in Judging the Strength of Evidence and Realistic Replication Expectations". Statistics in Biopharmaceutical Research Volume 12 (1), pages 43-45. DOI:10.1080/19466315.2020.1847182

- Sridhara R, Marchenko O, Jiang Q, Binkowitz B (2020). "Summary of American Association Biopharmaceutical Section's Virtual Discussion with Regulators on Type I error Considerations in Master Protocols with Common Control in Oncology Clinical Trials", *Biopharmaceutical Report Report*, Fall Issue, Volume 27 (3), pp. 39-40 : https://community.amstat.org/biop/biopharmreport
- Stallard, N., Todd, S., Parashar, D., Kimani, P. K., & Renfro, L. A. (2019). "On the need to adjust for multiplicity in confirmatory clinical trials with master protocols". *Annals of Oncology,* 30(4), pp.506. DOI: 10.1093/annonc/mdz038
- Wason, J. M., & Robertson, D. S. (2020). "Controlling type I error rates in multi-arm clinical trials: A case for the false discovery rate". *Pharmaceutical Statistics,* 20(1), pp. 109-116. DOI:10.1002/pst.2059
- Woodcock, J., and LaVange, L. M. (2017), "Master Protocols to Study Multiple Therapies, Multiple Diseases, or Both," *New England Journal of Medicine*, 377, pp. 62–70. DOI: 10.1056/NEJMra1510062.

Appendix

American Statistical Association Biopharmaceutical Section's

Virtual Discussion on: Type I error Considerations in Master Protocols with Common Control in Oncology Clinical Trials

Host: Statistical Methods in Oncology Scientific Working Group

October 8, 2020

8 am – 10 am EST (New York)

Agenda

Meeting Moderators:

Dr. Qi Jiang, Seagen, Co-chair of ASA BIOP Statistical Methods in Oncology Scientific Working Group

Dr. Olga Marchenko, Bayer, Co-chair of ASA BIOP Statistical Methods in Oncology Scientific Working Group

Dr. Rajeshwari Sridhara, Oncology Center of Excellence, FDA

1. 8 am – 8:15 am: Welcome and Introduction
    - Dr. Bruce Binkowitz, Shionogi, Chair of ASA Biopharmaceutical Section
    - Dr. Richard Pazdur and Dr. Rajeshwari Sridhara, Oncology Center of Excellence, FDA

2. 8:15 am – 9:15 am: Point-Counterpoint

    Academic representatives (20 minutes):
    - Prof. Martin Posch, Medical Statistics at the Medical University of Vienna
    - Prof. Mary Redman, Clinical Research Division, Fred Hutch

    Industry representatives (20 minutes):
    - Dr. Yevgen Tymofyeyev, Statistics and Decision Sciences, Janssen RD, J&J
    - Dr. Nicole Li,Biostatistics and Research Decision Sciences, Merck & Co

    Regulatory Agency representatives (20 mins):
    - Prof. Kit Roes, EMA
    - Members from FDA, EMA, HC, TGA, and PMDA

3. 9:15 am – 9:55 am: Panel Discussion: (a) situations where adjustment of type I error is not necessary, and (b) situations where adjustment of type I error is necessary
    - Dr. Marc Theoret (FDA), Dr. Yuan Li Shen (FDA), Dr. Thomas Gwise (FDA), Dr. Filip Josephson (EMA), Lorenzo Hess (SMC, Switzerland), Dr. Michael Coory (TGA, Australia), Andrew Raven (HC, Canada), Dr. Naoto Kotani (PMDA, Japan), Dr. Scott Berry (Berry Consultants), Dr. Richard Simon

4. 9:55 am – 10:00 am: Concluding remarks
    ASA and OCE
    - Dr. Olga Marchenko, Bayer, and Dr. Qi Jiang, Seagen, Co-chairs of ASA BIOP Statistical Methods in Oncology Scientific Working Group
    - Dr. Richard Pazdur and Dr. Rajeshwari Sridhara, Oncology Center of Excellence, FDA