ISSN: 2641-6336
DOI: 10.33552/ABBA.2023.05.000608
Annals of
Biostatistics & Biometric Applications
Iris Publishers

Review Article

Covariate Adjustment in Oncology Clinical Trials: Past, Present and Future

Jay Herson*

Senior Associate, Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland USA

***Corresponding author:** Jay Herson, Ph.D, Senior Associate, Biostatistics Johns Hopkins Bloomberg School of Public Health Baltimore, Maryland USA.

Received Date: March 31, 2023

Published Date: April 06, 2023

Abstract

Baseline covariate adjustment in randomized controlled trials (RCTs) in oncology with primary endpoint time to event has a long history. This paper reviews the past, present and future of this practice. We distinguish between stratification by baseline covariates which is used to gain precision in estimates and use of log linear models, such as proportional hazards regression, with or without stratification. The latter is done either to increase precision in estimates or to find which covariates may be considered prognostic factors. Knowledge of prognostic factors will allow us to define characteristics of patients who are likely to benefit from treatment and are a vital part of the new initiative for precision or personalized medicine. We provide insight into the effect of stratification on estimation of standard errors and size and power of statistical hypothesis tests and the use of minimization randomization as another method to create covariate balance between treatment groups. We close with current challenges In covariate adjustment in RCTs that use restricted mean survival time, win statistics, new estimands, allow treatment crossover and tissue-agnostic clinical trials.

Introduction

This paper deals with covariate adjustment in randomized controlled trials [RCT] with a time to event endpoint. These trials are usually the pivotal trial leading to a new drug application and occur at the end of a long clinical development program. The term survival time will be used throughout but time to event could also be progression-free survival time, time to dropout, time to adverse event, etc.

The covariates are patient data measured at baseline such as age, gender, stage of disease, white blood count, etc. Adjustment tends to provide a more precise estimate of the treatment effect. One obvious form of adjustment is stratified analysis. Although the latter will be mentioned throughout the emphasis will be placed on adjustment through log linear models. The log linear models can also be used to find covariates that are prognostic factors that might inform what types of patients can benefit from a particular intervention. This type of analysis is a vital part of the new initiative for precision or personalized medicine.

Past

There is a long history of baseline covariate adjustment in oncology. Originally analysis was based on landmark estimates e.g., three-year survival, five-year survival etc. This analysis did not consider dropouts or censored observations. The treatment effect was quantified by the odds ratio and the model used for adjustment was the logistic [1]. Later Kaplan-Meier life table methods were introduced [2]. These methods incorporated censored survival times. The Cox proportional hazards regression model used the Kaplan-Meier estimates to measure treatment effect by the hazard ratio [3]

Both the logistic and proportional hazards models allowed for assessing the relative importance of baseline variables on survival time. After several years of using these models to find prognostic factors it was found that the covariates associated with survival were strongly related to treatment. Predictive variables for, say, lung cancer varied depending on the treatments used. When ineffective treatments were used no prognostic factors were found because no patients had favorable survival. When highly effective treatments were used no prognostic factors emerged because all patients had favorable survival times regardless of baseline characteristics. It is in the middle ground that we found prognostic factors depending on the treatments used. Of course, the degree of censoring in the survival data can also influence our ability to identify prognostic factors.

Present

Proportional hazards regression methods and stratification are the principal methods of covariate adjustment today. Almost all pivotal RCTs use proportional hazards regression analysis for the survival endpoint and a stratified log rank test with stratification based on same variables used for randomization. One example is the Keynote 024 trial of pembrolizumab vs chemotherapy for PD-L1 positive non-small cell lung cancer, Here the stratification factors were ECOG performance status, histologic type and region of enrollment [4].

In the past, germline or somatic mutations might have been a covariate and / or stratification factor but today oncology trials are often conducted only for patients with specific mutations. An example would be the trial of the experimental treatment TAS-102 in patients with metastatic colon cancer and KRAS mutation. Within the RECOURSE RCT patients were stratified on KRAS type (wild type or mutant), time from initial diagnosis and metastasis and geographic region with a stratified log rank test for overall survival [5].

While details of stratified log rank test and covariate adjustment are usually well-specified in clinical trial protocols for final data analysis the details of if and how covariate adjustment will be handled in a planned interim analysis are often not specified. A common practice is use of a stratified log rank statistic without model-based covariate adjustment, but this must be pre-specified in the protocol and statistical analysis plan.

The U.S. Food and Drug Administration recently issued a non-binding draft guidance to industry on covariate adjustment [6]. This guidance calls attention to several items that might be overlooked by sponsors preparing new drug applications to the agency. The guidance distinguishes between conditional analysis of treatment effect (subgroup effects) and unconditional treatment effect (overall effect). Assuming treatment effect is measured by the hazard ratio, due to the principal of non-collapsibility the conditional treatment effects can differ from the unconditional treatment effects even when the treatment effect is identical across subgroups [7]. For this reason, sponsors must report subgroup-specific hazard ratios as well as overall hazard ratio even if the principal purpose of covariate adjustment was for increased precision rather than to identify prognostic factors. The FDA prefers that models that find treatment-covariate interactions not be submitted as primary analysis but as supporting analysis which can help the FDA write a drug label that prescribing physicians can understand.

Stratification is sometimes accompanied by minimization randomization. Under this scheme a random assignment can be overruled if such assignment would cause a trial-wide treatment imbalance in a covariate. The trial protocol would specify a threshold, t, say t=3. If a random assignment would cause an imbalance in any covariate greater than or equal to 3 then a biased coin randomization is performed where the other treatment is selected with high probability, say 0.90. The latter is included to preserve randomization. Although this method appears in oncology clinical trial protocols the regulatory attitude toward this procedure is still unclear. An early paper on minimization was by Simon and Pocock [8]. Taves provides an overview of this procedure [9].

Various authors have investigated the effect of stratification on baseline variables with or without proportional hazards adjustment in terms of estimation of standard errors, size, and power of statistical hypothesis tests [10-12] Groenwold, White and Donders at al have provided insight on handling of missing baseline covariates [13] and Herson has described procedures in the face of discovery of fraud in reported baseline covariates [14].

Future

In this section we look at emerging analytic methods and oncology trial types for which model adjustment for covariates have not yet been developed or are in their infancy. Stratified methodology without model adjustment would apply to all methods.

Researchers have raised questions about blindly applying the proportional hazards regression model when there is a question as to the applicability of the proportional hazards assumption.

Three new methods have been proposed for survival analysis that do not require this assumption. Restricted Mean Survival Time (RMST) considers the distance between survival curves occurring prior to a maximum time point. [15] The latter point is generally specified at a time post-treatment start where only a few censored patients remain in follow up. Karrison and Kocherginsky describe non-linear model adjustment for RMST through averaging over all covariates and show that this method can yield increased efficiency [16]. Pak et al [17] and Uno, Claggett and Tian et al [18] compare traditional survival analysis methods with RMST estimates for actual recent oncology clinical trials. While covariate adjustment can always be addressed by traditional stratification and there are many advantages of RMST over traditional methods a the authors are not presenting methods of model-based covariate adjustment. However, a preliminary analysis by RMST methods and a sensitivity analysis by RMST would yield sponsors and regulators important information on treatment effect although many would prefer covariate adjustment by non-linear models to be comparable with the proportional hazards tradition.

Another approach to survival analysis that does not require the proportional hazards assumption is found in a category becoming known as "win statistics". These methods compute wins and losses by considering all paired differences in survival time between patients within treatments. The wins are defined by the number of times the experimental treatment survival time exceeds a comparison with a control group survival time. Losses are when the difference is opposite. Censored observations are easily incorporated. The generalized pairwise comparison method considers the difference in wins and losses [19] while the win ratio method considers the ratio of wins to losses [20]. These methods can be applied on a stratified analysis, but no model covariate adjustment currently exists. Like RMST a these methods are today useful in sensitivity analysis.

Since the passage of the Twenty First Century Medicines Act by the U.S. Congress the FDA is interested in sponsors incorporating real world data (RWD) from electronic health records, insurance claims databases, disease registries, aggregator databases etc. [21]. The recent paper by Jamielita, Li and Burke et al indicates the complexity of analyzing RWD and taking baseline covariates into account [22].

The International Council on Harmonization (ICH) and FDA have issued their guidance E9 (R1) which indicates that sponsors must take intercurrent events like dropout, rescue medication etc. into account and this might require estimands other than the common treatment strategy estimand (also known as intent to treat) [23]. Ratitch, Bell, Mallinckrodt et al describe the new estimand landscape [24]. Some new estimands such as hypothetical and those depending on counterfactual assumptions will eventually require methods of covariate adjustment.

For ethical reasons many oncology clinical trial protocols today allow patients treated by the control group to crossover to the experimental treatment group in case of treatment failure or after unmasking at the end of the clinical trial. It is desired to analyze the entire survival time of all patients not just the time on their original treatment group. The counterfactual method of rank-preserving structural failure time model [25] is used in this analysis. Recently Korhonen, Zuber and Branson et al used this method to analyze survival time in the RECORD-1 trial of everolimus in metastatic renal cell carcinoma [26]. They assume that randomization is sufficient to ensure reasonable baseline balance in covariates, so no adjustment is attempted. In trials where the primary time to event endpoint is progression-free survival (PFS) traditional analysis with stratified log rank or proportional hazards regression may be sufficient. When FDA requests data on overall survival in the face of crossover the rank-preserving method without correction for covariates is considered useful as supportive information because the method depends on assumptions that are difficult to verify. Hence the intent-to-treat analysis with stratified log rank and proportional hazards regression analysis is still considered primary for overall survival.

We are now seeing the emergence of tissue-agnostic clinical trial designs. These trials do not treat patients of a single organ site but rather patients who were diagnosed with a certain mutation that appears in multiple organ sites. The FDA has issued a recent guidance for these trials [27] Seligson, Knepper and Ragg et al summarize recent trials for pembrolizumab, larotractinib and entrectinib across organ sites [28]. The ROAR trial for defratinib in patients with BRAF V600E-mutated rare cancers was an open label phase II trial [29]. A Bayesian hierarchical model is used with borrowing between histology types. It appears that currently the tissue-agnostic clinical trial designs are far from considering adjustment for covariates or even stratification. It will be interesting to see if oncologists are willing to put aside their well-known organ site prognostic baseline covariates when other prognostic factors emerge across organ sites. We must wonder if a lung cancer expert would accept prognostic factors that emerge as an average over various organ sites rather than the lung cancer prognostic factors that are well-accepted in the lung cancer community.

Conclusion

This paper has reviewed the history, present day and emerging issues of the important task of baseline data covariate adjustment in oncology pivotal RCTs with time to event endpoints. Stratified randomization, log linear models, non-parametric methods, or some combination of these. The common thread of covariate adjustment is to create balanced treatment comparisons and to use this information to identify patient types likely to benefit from treatment. Much research is needed to improve precision of estimates in the traditional oncology RCT but challenges for methodology development are needed to deal with methods that are used in lieu of proportional hazards regression, trials that use novel estimands, those that rely on real world data, trials that allow treatment crossover and the new tissue-agnostic clinical trials.

Acknowledgement

None.

Conflict of Interest

None.

References

1. Hosmer DW, Lemeshow S (2000) Applied Logistic Regression, New York: Wiley.
2. Kaplan EL, Meier P (1958) Nonparametric estimation from incomplete observations. Journal of the American Statistical Association 53: 457-480.
3. Cox DR (1972) Regression models and life tables (with discussion). Journal of the Royal Statistical Society B 34: 187-220.
4. Reck M, Rodriguez-Abreu D, Robinson AG, et al. (2016) Pembrolizumab vs chemotherapy for PD-L1-positive non-small-cell lung cancer. New England Journal of Medicine 375(19): 1823-1833.
5. Mayer RJ, Van Cutsen E, Falcone A, et al. (2015) Randomized trial of TAS-102 for refractory metastatic colorectal cancer. New England Journal of Medicine 372(20): 1909-1919.
6. U.S. Food and Drug Administration, Adjusting for Covariates in Randomized Clinical Trials for Drugs and Biological Products-Draft Guidance for Industry, 2021.

7. Agresti A (2002) Categorical Data Analysis, (2nd Edn)., Wiley.
8. Pocock SJ, Simon R (1976) Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. Biometrics 32:954-955.
9. Taves DR (2010) The use of minimization in clinical trials. Contemporary Clinical Trials 31(2): 180-184.
10. Green SB, Byar DP (1978) The effect of stratified randomization on size and power of statistical tests in clinical trials. Journal of Chronic Disease 31(6-7): 445-454.
11. Gail MH (1988) The effect of pooling across strata in perfectly balanced studies. Biometrics 44(1): 151-162.
12. Bugni F, Canay IA, Shaikh AM, et al. (2018) Inference under covariate-adaptive randomization. Journal of the American Statistical Association 113(524): 1784-1796.
13. Groenwold RHH, White IR, Donders AR, et al. (2012) Missing covariate data in clinical research: when and when not to use the missing-indicator method for analysis. Canadian Medical Association Journal 184(11): 1265-1269.
14. Herson J (2016) Strategies for dealing with fraud in clinical trials. International Journal of Clinical Oncology 21(1): 23-27.
15. Zhao L, Claggett B, Tian Lu, et al. (2016) On the restricted mean survival time curve in survival analysis. Biometrics 72(1): 215-221.
16. Karrison T, Kocherginsky M (2018) Restricted mean survival time: Does covariate adjustment improve precision in randomized clinical trials? Clinical Trials 15(2): 178-188.
17. Pak K, Uno H, Kim DH, et al. (2017) Interpretability of cancer clinical trial results using restricted mean survival time as an alternative to the hazard ratio. JAMA Oncology 3(12): 1692-1696.
18. Uno H, Claggett B, Tian L, et al. (2014) Moving beyond the hazard ratio in quantifying the between group differences in survival analysis. Journal of Clinical Oncology 32(22): 2380-2385.
19. Buyse M (2010) Generalized pairwise comparisons of prioritized outcomes in the two-sample problem. Statistics in Medicine 29(30): 3245-3247.
20. Pocock SJ, Ariti CA, Collier TJ, et al. (2012) The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. European Heart Journal 33(2):171-182.
21. Concato J, Corrigan-Curay J (2022) Real world evidence: where are we now? New England Journal of Medicine 386(18): 1680-1682.
22. Jamielita J, Li X, Burke T, et al. (2021) Augmenting real world data through modeling key clinical trial eligibility criteria: an example of patients with non-small cell lung cancer treated with pembrolizumab. JCO Clinical Cancer Informatics 5: 849-858.
23. US Food, Drug Administration. (2023) E9(R1) Statistical Principles for Clinical Trials: Addendum: Estimands and Sensitivity Analysis in Clinical Trials, Guidance for Industry 2021.
24. Ratitch B, Bell J, Mallinckrodt C, et al. (2020) Choosing estimands in clinical trials: putting the ICH E9(R1) into practice. Therapeutic innovations and Regulatory Science 54(2): 324-341.
25. Robins JM, Tsiatis AA (1991) Correcting for non-compliance in randomized trials using rank- preserving structural failure time models. Communications in Statistics-Theory and Methods 20(8): 2609-2631.
26. Korhonen P, Zuber E, Branson M, et al. (2012) Correcting overall survival for the impact of crossover via a rank-preserving failure time (RPSFT) model in the RECORD-1 trial of everolimus in metastatic renal cell carcinoma. Journal of Biopharmaceutical Statistics 22(6): 1258-1271.
27. US Food, Drug Administration, Tissue Agnostic Drug Development in Oncology, Draft Guidance for Industry, 2022.
28. Seligson ND, Knepper TC, Ragg S, et al. (2021) Developing drugs for tissue-agnostic indications: a paradigm shift in leveraging cancer biology for precision medicine. Clinical Pharmacology and Therapeutics 109(2): 334-342.
29. Wen PY, Stein A, Van den Bent M, et al. (2022) Dabrafenib plus trametinib in patients with BRAF V600E-mutant low-grade and high-grade glioma - ROAR: a multicentre, open-label, single-arm phase 2, basket trial. The Lancet Oncology 23(1): 53-64.