

Application and Challenges of Crossover Trial Designs in the Fields of Treatment and Prevention

Junling Feng, Baoning Qi*

Department of Public Health, Shaanxi University of Chinese Medicine, Xianyang, Shaanxi, 712046, China

**Corresponding author*

Keywords: Crossover; tumor treatment; prevention

Abstract: Crossover designs, as a methodological approach in clinical trials, offer advantages and potential across both therapeutic and preventive domains, despite facing unique challenges. By enabling both intra-group and inter-group comparisons, this design effectively reduces the required number of participants and trial costs while mitigating the impact of sequence effects. Consequently, it provides an efficient framework for the assessment of drug efficacy and vaccine effectiveness. In therapeutic contexts, crossover trials enhance precision medicine initiatives and support the development of personalized treatment strategies. In preventive settings, they strengthen the validation of long-term vaccine efficacy. Despite their diverse applications, crossover trials share a common objective: to improve medical outcomes and protect public health. Looking ahead, with advancements in technologies such as big data analytics and artificial intelligence integration, crossover trials are poised to uncover even broader opportunities and profound impacts within both therapeutic and preventive fields.

1. Introduction

Phase III Randomized Controlled Trials (RCTs) represent the gold standard for assessing the efficacy and safety of medications, where subjects are randomly assigned to different treatment groups (Group A and Group B) to ensure objectivity and reliability of study outcomes^[1]. Crossover trial design, a specific type of RCTs, dynamically adjusts the treatment allocation based on the participants' response during the trial, enabling continuous monitoring and evaluation of drug efficacy. This design not only effectively mitigates potential confounding factors present in parallel-group RCTs but also significantly reduces sample size requirements while maintaining comparable precision in effect estimation and control of Type I error rates. It is particularly suited for managing symptoms of chronic or stable conditions such as multiple sclerosis, rheumatoid arthritis, and chronic obstructive pulmonary disease. However, amidst the global spread of the COVID-19 pandemic, crossover trial designs have demonstrated new potential applications in preventive medicine, notably in vaccine development. The Food and Drug Administration has issued new guidelines for COVID-19 vaccines (e.g., Pfizer-BioNTech, AstraZeneca), allowing for expedited validation and dissemination of vaccine efficacy through interim authorizations that bridge immunity to populations not initially included in trials. Flexible trial designs facilitate this

process.

Crossover trial designs can be further classified into systematic and non-systematic crossover based on the differing strategies employed for crossover. Systematic crossover is guided by explicit rules for group switching, whereas non-systematic crossover relies on subjective judgments by the investigators. However, crossover trials still face significant challenges in clinical data analysis. Factors such as rapid disease progression, repeated measurement of outcome events, or prolonged study durations, prevalent in fields like oncology clinical trials and vaccine-induced protection studies, can complicate the accurate assessment of drug/vaccine efficacy due to the confounding effects of carryover and residual effects. To address these challenges, various statistical methodologies (including generalized linear models, mixed-effects models, Cox proportional hazards models with time-dependent covariates) have been proposed to reduce bias and accurately estimate treatment effects. Nevertheless, the selection of these methods is based on certain assumptions and can critically impact the final analysis of treatment efficacy.

Therefore, there is a pressing need to delve deeper into the statistical methodology choices for crossover designs in both treatment and prevention domains, aiming to identify a universal approach that can accurately evaluate not only oncological interventions but also long-term vaccine efficacy. The choice of methodology influences both patient treatment plans and the implementation of public health policies, underscoring its paramount importance.

2. Statistical Methodology Research

2.1 Assessment Methods for the Efficacy of Treatment Drugs

In the clinical trial domains of oncology treatment and medical device usage, researchers have developed specialized statistical methods tailored for analyzing crossover trial designs, addressing a diverse range of study endpoints including binary, ordinal, and continuous outcomes. For continuous endpoints that follow a normal distribution, Bellavance et al. proposed a modified F-test approximation method, which effectively addresses the issue of intra-subject correlation arising from repeated measurements. The adequacy of this method in controlling type I error was validated through simulation of a 3x3 crossover trial^[2]. Furthermore, since each subject receives multiple treatments and may exhibit sequence or carryover effects, models capable of handling these complex factors are necessary. Hulst et al. integrated the use of paired-sample t-tests, Mann-Whitney U test and generalized linear mixed models to thoroughly evaluate the data for carryover and period effects.^[3] Currently, general linear models and linear mixed models have garnered significant attention in the research field of treatment under crossover designs due to their simplicity in model construction, execution, and result interpretation, making them the predominant statistical methods.

Wei Wang et al. evaluated the estimation performance of general linear model (GLM) in estimating treatment effects, controlling type I error, and testing for treatment effects through simulation of a 2x2 repeated measures crossover trial comparing a novel test contact lens with a control lens^[4]. By comparing s sequence p period crossover clinical trial, assuming there are subjects within sequence groups in the crossover trial, let represent the response observed for the k subject in period j of sequence group i , where it follows a continuous and normal distribution.

$$y_{ijk} = \mu + \pi_j + \tau_{d[i,j]} + \lambda_{d[i,j-1]} + \epsilon_{ijk}$$

Wherein, μ represents the intercept, π_j represents the period-related effect associated with the period j , $\tau_{d[i,j]}$ denotes the direct treatment effect related to the treatment applied in the period j of the sequence i , $\lambda_{d[i,j-1]}$ represents the carryover effect from the previous period of the sequence i ,

and ϵ_{ijk} is the random error with zero mean and variance.

In 3x3 crossover trials with a large number of sequences, GLM that incorporate carryover and period effects can offer significantly higher statistical power than single-factor ANOVA methods. This model assumes that differences in residual effects are proportional to treatment effects and is recommended for estimating treatment effects in 2x2 crossover designs only when residual effects are minimal. Furthermore, this model considers only 2x2 crossover trials involving first-order carryover effects, and results beyond the scope of these assumptions have not been further analyzed.

To evaluate the accuracy and efficiency of this weighted average method in estimating treatment effects, a 2x2 crossover design was simulated[5]. To evaluate the accuracy and efficiency of this weighted average method in estimating treatment effects, a 2x2 crossover design was simulated. For patient i assigned to group g , let $Y_{iz}^{(g)}$ denote the subject's response in period z , which is represented by the following random effects linear mixed model,

$$Y_{iz}^{(g)} = \mu_i^{(g)} + \eta_{BA}X_{iz}^{(g)} + \gamma Z_{iz}^{(g)} + \lambda_A I_{\{g=1\}} Z_{iz}^{(g)} + \lambda_B (1 - I_{\{g=1\}}) Z_{iz}^{(g)} + \epsilon_{iz}^{(g)}$$

Where $\mu_i^{(g)}$ represents the random effect of i patient in group g , $X_{iz}^{(g)}$ represents the covariate of the treatment group B, $Z_{iz}^{(g)}=1$ indicates that the i patient in group g receives treatment B in period z , otherwise it is 0; $I_{\{g=1\}}$ is an indicator function of group A; $\epsilon_{iz}^{(g)}$ represents the random error. η_{BA} and γ represent the difference between treatment B and treatment A, as well as the difference between period 2 and period 1. Furthermore, λ_A and λ_B represent the residual effect caused by treatment A and B.

However, the model has yet to consider the validity of its assumptions, particularly the assumption that the random effects linear mixed model has independently and identically distributed random errors, which can be difficult to uphold in real-world scenarios, potentially leading to errors in model estimation results. Secondly, the model only considers inter-treatment period effects and random effects, neglecting other potential factors that may influence the outcomes, such as individual differences and measurement errors.

Moses Mwangi et al. proposed the use of a piecewise linear mixed-effects (PLME) model to evaluate drug effects. This model offers flexibility in handling drug responses across different phases of the period, making it suitable for capturing changes in response patterns over time.[6]

$$Y_{ij} = \beta_0 + \beta_1 t_j + \beta_2 P_j + \beta_3 T_{ij} + \beta_4 P_j t_j + \beta_5 T_{ij} t_j + \beta_6 P_j T_{ij} + \beta_7 P_j T_{ij} t_j + b_{0i} + b_{1i} t_j (1 - P_j) + b_{i2} t_j P_j + \epsilon_{ij}$$

Among them, the time of the i participant was allocated to treatment A when $T_{ij} = 0$, and if assigned to treatment B when $T_{ij} = 1$. P_j represents the j period. β_6 and β_7 represent fixed effects resulting from the interaction of periodic and periodic treatment. b_{0i} is the random intercept of the i participant, b_{1i} is the slope of the i participant before the random change in the first period $P_j = 0$, b_{i2} is the slope of the i participant after the random change in the second period $P_j = 1$, and ϵ_{ij} is the random error measured by the i participant assigned to treatment T , during period P_j , and time period T_j . β_0 is total intercept term. β_1 and β_2 represent the fixed effects of periodic treatment and periodic treatment interactions.

This model further describes the impact of time, treatment, and period on the estimation of treatment effects. By setting t_4 as the crossover time, it divides the treatment groups $T_{ij} = 0$ into two stages, before and after the transition, allowing for separate calculations of subject responses within each stage.

$$Y_{ij} = t_j^* = \begin{cases} \beta_0 + \beta_1 t_j + b_{10} + b_{i1} t_j + \epsilon_{ij} & t_j \leq t_4 \\ (\beta_0 - t_4 \beta_3) + (\beta_1 + \beta_3) t_j^* + b_{i2} t_j^* + \epsilon_{ij} & t_j > t_4 \end{cases}$$

The control group $T_{ij} = 1$ are analyzed separately for the two stages before and after the crossover.

$$Y_{ij} = \begin{cases} (\beta_0 + \beta_4) + (\beta_1 + \beta_5) t_j + b_0 + b_{i1} t_j + \epsilon_{ij} & t_j \leq t_4 \\ (\beta_0 + \beta_4) - t_4(\beta_3 + \beta_7) + (\beta_1 + \beta_3 + \beta_5 + \beta_7) t_j^* + b_2 t_j^* + \epsilon_{ij} & t_j > t_4. \end{cases}$$

The PLME model, characterized by its segmented linear structure, enables the independent modeling of linear variations in the response variable across diverse time periods. By incorporating random effects to account for individual heterogeneity, it significantly enhances the accuracy of estimation. This model exhibits flexible adaptability to various experimental designs, particularly excelling in the analysis of changes in treatment effects before and after transitions. Nevertheless, further research and optimization are indispensable for assessing its goodness of fit, predictive capabilities, and applicability in multi-period and multi-treatment group scenarios. In randomized crossover trials, PLME holds promise in estimating nonlinear trends in treatment effects while effectively controlling for residual errors. Nonetheless, future studies should focus on enhancing model optimization, validation procedures, and cross-design applications.

2.2 Methods for Evaluating Vaccine Efficacy

Methodological research in cross-over trials within the field of prevention primarily focuses on evaluating the effectiveness, safety, and interactions among various preventive measures, such as vaccination and public health interventions. Notably, during large-scale outbreaks, the assessment of vaccine protection efficacy is conducted by comparing the risk after vaccination in the placebo group with the risk at the start of the trial in the vaccinated group.[7,8]The Poisson distribution-based method for estimating vaccine efficacy (VE) is a commonly used statistical approach in the assessment of vaccine effectiveness, particularly in clinical trials designed with a fixed number of events. Since VE assessment often relies on the observed number of disease occurrences within a certain period, the Poisson distribution, which is suitable for describing the number of random events occurring within a fixed time or space, is particularly useful in VE evaluations.[9]Furthermore, due to the potential emergence of multiple novel strains or variants during large-scale outbreaks, the transmission speed of the virus accelerates, and vaccines developed based on the original virus strain may exhibit reduced efficacy against the new strains compared to the original strain. Dean Follmann et al. proposed that during the conduct of phase III clinical trials for COVID-19 vaccines, it is necessary to establish the initial efficacy of the vaccine against the original strain before crossover. By comparing the incidence of COVID-19 cases infected with the Delta variant in the originally vaccinated group after crossover, the VE can be estimated. Specifically, a higher incidence rate in the originally vaccinated group after crossover indicates a weakened VE.[10]

When not considering strain variation,

$$VE_K = 1 - \frac{\omega_{1K}}{\theta_K}$$

Where ω_{AK} is the mean value of x_{AK} (the number of cases in group A in period K) approximating Poisson distribution, and $K=1$ and 2 are the stages before and after group transfer. θ_K is the expected number of placebo cases in phase K. Before switching groups (in case of strain change,

$$VE_1 = 1 - \frac{X_{11}}{X_{01}}$$

After switching groups (in case of strain change), X_{AK} indicates that the number of cases in group A followed Poisson distribution in period K,

$$VE_2 = 1 - \frac{X_{11} X_{12}}{X_{01} X_{02}}$$

When strain variations are not considered, this approach can be extended to the modified Poisson method for estimating the overall vaccine efficacy, providing estimates of relative risk and vaccine efficacy for specific time periods. However, in the real world, it is crucial to pay attention to issues such as data integrity and accuracy, sample size considerations, testing of model assumptions, and comparison of multiple methods. To address these challenges, Dean Follmann further proposes a modified time-dependent Cox model that accommodates staggered enrollment and crossover, dropouts, covariate adjustments, changes in placebo event rates over calendar time, stratification by study centers, and smooth changes in strain-specific VE over time to enable more accurate VE estimation.[10,11].

When considering only the crossover design,

$$h_s(t) = h_0(t) \exp \left[Z_i(t) \left\{ \beta_{0s} + \beta_{1s} (t - \tau_i^{(v)}) \right\} \right] I(\tau_i^{(e)} > t)$$

When considering changes in virus strains after crossover,

$$h_s(t) = h_0(t) P(t, s) \exp \left[Z_i(t) \left\{ \beta_{0s} + \beta_{1s} (t - \tau_i^{(v)}) \right\} \right] I(\tau_i^{(e)} > t)$$

Where, t is the study start time, $\tau_i^{(e)} \tau_i^{(v)}$ is the enrollment and vaccination time of subject i , $h_0(t)$ is the risk of infection of any strain and $P(t, s)$ is the true proportion of strains in the surveillance cohort. $Z_i(t)$ is the indicator of whether subject i was vaccinated at the time.

The time-dependent Cox model flexibly estimates vaccine efficacy as virus strains and trial designs evolve, taking into account factors such as seasonality, covariates, crossover populations, and crossover rates. However, this model primarily focuses on a unidirectional crossover pattern, where participants in the placebo group receive the trial vaccine after a delay. The model has not yet incorporated the impact of factors such as increased vaccine doses, spatial heterogeneity, heterogeneity in vaccine protection, and individual immune responses on VE estimation. In summary, a detailed examination of the statistical methodologies used in crossover designs in cancer treatment and prevention reveals that different methods are often employed to assess the efficacy of drugs or vaccines across various fields. Nonetheless, crossover designs share the common objective of evaluating therapeutic outcomes, whether in treatment or prevention.

3. Convergence in Divergent Paths

3.1 Differences in Crossover Designs Applied in Therapeutic and Preventive Domains

For conditions with long treatment courses or chronic diseases, selective crossover designs are often adopted to adjust trial groups, but this can introduce confounding from the combined effects of the investigational drug and placebo. Additionally, due to the heterogeneity of tumor diseases and the diversity of treatment regimens, participants are typically individuals who are already ill or at high risk, and their physical condition and disease status may significantly impact trial outcomes. Therefore, when applying crossover designs, factors such as disease stage and underlying diseases

must be taken into consideration.[12]In clinical data analysis, the evaluation of drug efficacy often involves the use of complex statistical models such as the Cox proportional hazards model, generalized linear models, and mixed-effects models to handle survival data and time-dependent variables. These models necessitate consideration of factors like lag effects, inter-patient heterogeneity, and drug interactions. Evaluation metrics commonly include disease remission rates, cure rates, and survival rates[13]. Helgestad conducted a non-blinded, multi-phase, cluster-randomized crossover trial and employed logistic regression with cluster-robust standard errors to conduct a sensitivity analysis on the primary outcome, examining the promotional effect of breast cancer screening on cervical cancer screening[14]. Furthermore, as the treatment process for diseases is often lengthy, and long-term efficacy and safety of different treatment regimens need to be evaluated, the trial duration is typically prolonged, resulting in higher research costs, including drug costs, patient follow-up costs, data collection and analysis costs, among others.

In contrast, vaccine trials often adopt a unidirectional crossover design, where participants are typically healthy individuals or those at low risk, with relatively good physical conditions that introduce less interference to trial outcomes. The implementation of vaccine trials is standardized and straightforward, not involving complex disease management.[15,16]

The evaluation of long-term vaccine effectiveness focuses on assessing whether the vaccine can generate sufficient immune protection. Some scholars propose extensions to non-proportional hazards models and time-dependent risk models. Additionally, factors such as the long-term effects after vaccination, differences among different populations, and the synergistic effects of vaccines with other public health measures need to be considered. Evaluation metrics include immunological indicators like seroconversion rate and duration of antibody. Moreover, the effects of preventive measures designed with crossover designs (e.g., vaccination) can often be observed within a relatively short period, resulting in shorter trial durations and lower costs, with minimal impact from washout periods[17-21].

3.2 Similarities in Crossover Designs across Therapeutic and Preventive Domains

Crossover designs exhibit distinctive flexibility and value in both therapeutic and preventive medical applications. Firstly, from an ethical and participant rights standpoint, these designs rigorously comply with ethical standards, ensuring the protection of participant rights, enhancing adherence to treatment or prevention protocols, thus reducing disease risk, improving public health, and offering additional options for clinical trial participants, which ultimately fosters greater patient engagement and satisfaction. Secondly, concerning the selection of statistical methods, although the therapeutic and preventive domains have distinct emphases, there is a notable degree of complementarity between the methodologies. Specifically, assessing the efficacy of drugs and vaccines requires careful consideration of treatment sequence effects and period effects on outcomes. The Inverse Probability of Censoring Weighting (IPCW) method enhances the robustness of statistical inferences by mitigating the impact of crossover behavior on statistical analysis through weighted adjustments, allowing researchers to more accurately estimate the effects of therapeutic or preventive interventions and reduce statistical errors due to missing data and confounding biases.

However, crossover designs encounter challenges such as extended follow-up periods and difficulties in maintaining quality control, irrespective of their application in therapeutic or preventive contexts. To overcome these challenges, it is essential to enhance clinical trial management, optimize follow-up procedures, and continuously investigate and implement more advanced statistical methods to ensure the precision and reliability of trial outcomes. In conclusion, the utilization of crossover designs in the medical field not only enhances therapeutic and

preventive efficacy but also advances research methodologies and protects participant rights. This design approach represents a significant contribution to medical research.

4. Conclusions

The widespread application of crossover designs in the medical field is not only evident in their meticulous data comparison and analysis but also in their ability to explore therapeutic and preventive effects across multiple dimensions and factors, providing a solid empirical foundation for medical advancements. By implementing different intervention measures within the same population and comparing them at distinct time points or under varying conditions, crossover designs effectively mitigate biases arising from external factors such as individual differences and temporal variations. Consequently, the research outcomes are more reliable, accurately reflecting the true effects of interventions like drugs or vaccines.

Although crossover designs exhibit differences in research design and objectives, design types and complexity, statistical method selection, subject selection, and trial costs between oncology treatment and prevention fields, they enhance research quality and efficiency in both therapeutic and preventive domains through various experimental designs (e.g., selective crossover design, cluster crossover design, one-way crossover design) and statistical approaches (e.g., generalized linear models, Cox regression models). Furthermore, they facilitate the development of personalized medicine, multidisciplinary collaboration, and medical innovation, ultimately improving medical outcomes and safeguarding public health.

Overall, this methodology is well-suited for estimating the effectiveness of drugs and vaccines, offering flexibility in providing more personalized treatment regimens tailored to patients' varying conditions, ages, and physical constitutions. It also enables the formulation of more targeted preventive measures addressing risk factors and lifestyle habits across different populations. By conducting multiple measurements on the same patient, crossover designs reduce the need for large sample sizes, thereby lowering research costs. Therefore, crossover design clinical trial methodologies converge in their ultimate goals in both tumor treatment and disease prevention, with their research findings providing crucial evidence for the formulation of scientific and rational prevention and treatment strategies. This, in turn, effectively reduces the burden of disease and enhances public health. In the future, with the continuous integration of technologies such as big data and artificial intelligence, crossover designs are poised to exhibit even broader application prospects and profound influence in the fields of treatment and prevention.

References

- [1] Akova M, Unal S. A randomized, double-blind, placebo-controlled phase III clinical trial to evaluate the efficacy and safety of SARS-CoV-2 vaccine (inactivated, Vero cell): a structured summary of a study protocol for a randomised controlled trial[J]. *Trials* 2021,22(1):276.
- [2] Bellavance F, Tardif S, Stephens MA. Tests for the analysis of variance of crossover designs with correlated errors[J]. *Biometrics* 1996,52(2):607-612.
- [3] van Hulst AM, van den Akker E, Verwaaijen EJ, et al. Hydrocortisone to reduce dexamethasone-induced neurobehavioral side-effects in children with acute lymphoblastic leukaemia-results of a double-blind, randomised controlled trial with cross-over design[J]. *Eur J Cancer* 2023,187:124-133.
- [4] Wang W, Cong N, Chen T, et al. A note on misspecification in general linear models with correlated errors for the analysis of crossover clinical trials[J]. *PLoS One* 2019,14(3):e213436.
- [5] Lin CD, Lui KJ. A note on point estimation and interval estimation of the relative treatment effect under a simple crossover design[J]. *Pharm Stat* 2022, 21(2):386-394.
- [6] Mwangi M, Molenberghs G, Njagi EN, et al. Pairwise fitting of piecewise mixed models for the joint modeling of multivariate longitudinal outcomes, in a randomized crossover trial[J]. *Biom J* 2024,66(2):e2200333.
- [7] Poland GA, Ovsyannikova IG, Kennedy RB. SARS-CoV-2 immunity: review and applications to phase 3 vaccine

- candidates[J]. *Lancet* 2020,396(10262):1595-1606.
- [8] Choe PG, Kang CK, Suh HJ, et al. Waning Antibody Responses in Asymptomatic and Symptomatic SARS-CoV-2 Infection[J]. *Emerg Infect Dis* 2021,27(1):327-329.
- [9] Follmann D, Fintzi J, Fay MP, et al. A Deferred-Vaccination Design to Assess Durability of COVID-19 Vaccine Effect After the Placebo Group Is Vaccinated[J]. *Ann Intern Med* 2021,174(8):1118-1125.
- [10] Follmann D, Mateja A, Fay MP, et al. Durability of Protection Against COVID-19 Through the Delta Surge for the NVX-CoV2373 Vaccine[J]. *Clin Infect Dis* 2024,79(1):78-85.
- [11] Follmann D, Fay M, Magaret C. Estimation of vaccine efficacy for variants that emerge after the placebo group is vaccinated[J]. *Stat Med* 2022,41(16):3076-3089.
- [12] Hattswell A, Freemantle N, Baio G, et al. Summarising salient information on historical controls: A structured assessment of validity and comparability across studies[J]. *Clin Trials* 2020,17(6):607-616.
- [13] Nathan SD, Johri S, Joly JM, et al. Survival analysis from the INCREASE study in PH-ILD: evaluating the impact of treatment crossover on overall mortality[J]. *Thorax* 2024,79(4):301-306.
- [14] Helgestad A, Larsen MB, Njor S, et al. Increasing coverage in cervical and colorectal cancer screening by leveraging attendance at breast cancer screening: A cluster-randomised, crossover trial[J]. *PLoS Med* 2024,21(8):e1004431.
- [15] Tsiatis AA, Davidian M. Estimating vaccine efficacy over time after a randomized study is unblinded[J]. *Biometrics* 2022,78(3):825-838.
- [16] Follmann D, Fintzi J, Fay MP, et al. Assessing Durability of Vaccine Effect Following Blinded Crossover in COVID-19 Vaccine Efficacy Trials[J]. *medRxiv*, 2020.
- [17] Fintzi J, Follmann D. Assessing vaccine durability in randomized trials following placebo crossover[J]. *Stat Med* 2021,40(27):5983-6007.
- [18] Lin DY, Zeng D, Gilbert PB. Evaluating the Long-Term Efficacy of COVID-19 Vaccines[J]. *medRxiv*, 2021.
- [19] Lin DY, Zeng D, Gu Y, et al. Reliably Assessing Duration of Protection for Coronavirus Disease 2019 Vaccines[J]. *J Infect Dis* 2022,226(11):1863-1866.
- [20] Nikas A, Ahmed H, Zarnitsyna VI. Competing Heterogeneities in Vaccine Effectiveness Estimation[J]. *ArXiv*, 2023.
- [21] Nikas A, Ahmed H, Zarnitsyna VI. Estimating Waning of Vaccine Effectiveness: A Simulation Study[J]. *Clin Infect Dis* 2023,76(3):479-486.