Model-Robust Inference for Clinical Trials that Improve Precision by Stratified Randomization and Adjustment for Additional Baseline Variables

Bingkai Wang[1], Ryoko Susukida[2], Ramin Mojtabai[2], Masoumeh Amin-Esmaeili[2,3], Michael Rosenblum[1]

[1] Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health
[2] Department of Mental Health, Johns Hopkins Bloomberg School of Public Health
[3] Iranian National Center for Addiction Studies (INCAS), Tehran University of Medical Sciences

For clinical trials using stratified randomization, we can do the analysis potentially better.

“clinical trials”

For participant $i = 1, \ldots, n$, we observe

- Y_i is the outcome variable, which can be continuous, binary or time-to-event,
- A_i is a binary treatment indicator,
- X_i is a vector of baseline variables.

We use the Neyman-Rubin causal model and assume

$$Y_i = A_i Y_i(1) + (1 - A_i) Y_i(0),$$

where $Y_i(1)$ and $Y_i(0)$ are potential outcomes.

Our goal is to estimate a population parameter, for example, the average treatment effect (ATE)

$$\Delta = E[Y_i(1) - Y_i(0)].$$

“stratified randomization”

“stratified randomization” refers to stratified permuted block randomization, which is used by 70% of trials in top medical journals in 2014 (Lin et al., 2015).

For example, suppose that there are 2 strata {Female, Male} and the block size is 4.

Male	1	0	1	0	0	0	1	1	...
Female	1	1	0	0	0	1	1	0	...

“stratified randomization”

Stratified randomization is different from simple randomization:

1. Stratified randomization can ensure treatment balance within each stratum: $\sum_{i=1}^{n} A_i \approx n - \sum_{i=1}^{n} A_i$.
2. Treatment of different participants are not longer independent.

However, people usually ignore this difference in statistical analysis.

According to a survey by Kahan and Morris (2012), only 35% trials in top medical journals in 2010 adjusted for strata in their analysis.

“the analysis”

For binary or continuous outcomes, “the analysis” refers to any M-estimator of Δ, which includes:

- the ANCOVA estimator for continuous outcomes,
- the standardized logistic regression estimator for binary outcomes,
- doubly-robust weighted-least-square estimator (involving missing outcomes under the missing at random assumption),
- augmented inverse-probability-weighted (AIPW) estimator,
- mixed-effects model for repeated measures (MMRM),
- targeted maximum likelihood estimator (TMLE).

For time-to-event outcomes, “the analysis” refers to the Kaplan Meier estimator of survival function $P(Y_i(a) > t)$ for $a = 0, 1$ and $t \in [0, \tau]$.

“potentially better”

“Potentially better” means potentially smaller asymptotic variance.

Theorem 1

For any M-estimator $\hat{\Delta}$ such that $\sqrt{n}(\hat{\Delta} - \Delta) \xrightarrow{d} N(0, V_{\mathrm{simple}})$ under simple randomization, we have $\sqrt{n}(\hat{\Delta} - \Delta) \xrightarrow{d} N(0, V_{\mathrm{strat}})$ under stratified randomization with

$$V_{\mathrm{strat}} \leq V_{\mathrm{simple}}.$$

For the Kaplan Meier estimator, we establish the same result for stochastic process.

For clinical trials using stratified randomization, we can do the analysis potentially better.

$$\Downarrow$$

For a wide class of estimators, stratified randomization may lead to smaller variance than simple randomization.

$$\Downarrow$$

Use the correct variance V_{strat} when doing inference to avoid being conservative.

Related work

- Shao et al. (2010); Shao and Yu (2013) proved the validity of the two-sample t-test under the biased-coin design assuming generalized linear model.
- Ma et al. (2015, 2018) assumed a linear model and derived the asymptotic distribution of the test statistic of ATE for the ANCOVA estimator and a class of covariate-adaptive designs.
- **Bugni et al. (2018) established the asymptotic theory of the unadjusted estimator and the ANCOVA estimator (adjusting for strata only) of ATE for a wide range of covariate-adaptive designs.**
- Ye and Shao (2019) derived asymptotics for log-rank and score tests in survival analysis under covariate adaptive randomization.
- Li and Ding (2019) established the asymptotic theory for the ANCOVA estimator under covariate-adaptive randomization in the randomization inference framework.

Main contribution

1. We generalize the result by Bugni et al. (2018) to handle various outcome types, repeated-measured outcomes, missing data and covariate adjustment.
2. We prove the asymptotic result for statistical processes under stratified randomization and apply it to the Kaplan-Meier estimator.
3. We give consistent variance estimators for V_{strat} and R functions for implementation.
4. The above results also hold for the biased-coin covariate-adaptive design.

How much variance reduction can we have?

Example 1: using the correct variance formula V_{strat}

CTN44 is a study evaluating internet-delivered treatment for substance abuse.

- Outcome: time to abstinence.
- Treatment: Therapeutic Education System versus Treatment as usual.
- Stratification: patient's primary substance of abuse and abstinence status at baseline (4 strata).

Visit	1	2	3	4	5	6	7	8
Survival probability	0.58	0.53	0.47	0.40	0.39	0.33	0.30	0.27
Variance reduction $(1 - V_{strat}/V_{simple})$	11%	12%	11%	9%	7%	4%	3%	2%

Example 2: adjusting for additional baseline variables

CTN03, CTN30 and CTN44 are studies of treatment of substance use disorder using stratified randomization.

- CTN03 has binary outcomes and CTN30 and CTN44 have continuous outcomes, all being measures of treatment success.
- Each study has ∼5 baseline variables.

Study	Number of Strata	Unadjusted estimator (95% CI)	Adjusted estimator (95% CI)	Proportional variance reduction
CTN03	3	-0.11(-0.21, -0.01)	-0.10(-0.19, -0.02)	**35%**
CTN30	4	0.02(-0.02, 0.05)	0.01(-0.02, 0.04)	**17%**
CTN44	4	-0.09(-0.14, -0.03)	-0.09(-0.14, -0.03)	**2%**

Take-away message 1:

For clinical trials using stratified randomization, we may overestimate the variance of an estimator if we ignore the difference between stratified randomization and simple randomization. (Rosenblum and Wang, 2019)

Suggestion:

Do statistical inference using the correct variance (based on V_{strat} instead of V_{simple}).

Take-away message 2:

Adjusting for a set of preplanned baseline variables may lead to substantial variance reduction.

Limitation:

- It only works for phase 2 or 3 trials with large sample size.
- It requires that the number of subjects in each stratum is not small.

Thank you!

The slides are available at
https://bingkaiwang.com.

The paper is available at
https://arxiv.org/abs/1910.13954.

The R code is available at
https://github.com/BingkaiWang/covariate-adaptive.

Acknowledgement

This project was supported by a research award from Arnold Ventures. The content is solely the responsibility of the authors and does not necessarily represent the official views of Arnold Ventures. The information reported here results from secondary analyses of data from clinical trials conducted by the National Institute on Drug Abuse (NIDA). Specifically, data from NIDA–CTN-0003 (Suboxone (Buprenorphine/Naloxone) Taper: A Comparison of Two Schedules), NIDA-CTN-0030 (Prescription Opioid Addiction Treatment Study) and NIDA-CTN-0044 (Web-delivery of Evidence-Based, Psychosocial Treatment for Substance Use Disorders) were included. NIDA databases and information are available at (https://datashare.nida.nih.gov).

Bugni, F. A., I. A. Canay, and A. M. Shaikh (2018). Inference under covariate-adaptive randomization. *Journal of the American Statistical Association 113*(524), 1784–1796.

Kahan, B. C. and T. P. Morris (2012). Improper analysis of trials randomised using stratified blocks or minimisation. *Statistics in Medicine 31*(4), 328–340.

Li, X. and P. Ding (2019). Rerandomization and regression adjustment. *arXiv https://arxiv.org/abs/1906.11291*.

Lin, Y., M. Zhu, and Z. Su (2015). The pursuit of balance: An overview of covariate-adaptive randomization techniques in clinical trials. *Contemporary Clinical Trials 45*, 21 – 25. 10th Anniversary Special Issue.

Ma, W., F. Hu, and L. Zhang (2015). Testing hypotheses of covariate-adaptive randomized clinical trials. *Journal of the American Statistical Association 110*(510), 669–680.

Ma, W., Y. Qin, Y. Li, and F. Hu (2018). Statistical inference of covariate-adjusted randomized experiments. *arXiv https://arxiv.org/abs/1807.09678*.

Rosenblum, M. and B. Wang (2019, 04). The Critical Role of Statistical Analyses in Maximizing Power Gains From Covariate-Adaptive Trial Designs. *JAMA Network Open 2*(4), e190789–e190789.

Shao, J. and X. Yu (2013). Validity of tests under covariate-adaptive biased coin randomization and generalized linear models. *Biometrics 69*(4), 960–969.

Shao, J., X. Yu, and B. Zhong (2010). A theory for testing hypotheses under covariate-adaptive randomization. *Biometrika 97*(2), 347–360.

Ye, T. and J. Shao (2019). Validity and robustness of tests in survival analysis under covariate-adaptive randomization. *arXiv https://arxiv.org/abs/1811.07232*.