Regulatory perspectives on covariate adjustment in randomized trials

Robin Ristl

Center for Medical Data Science, Medical University of Vienna

Disclaimer

This presentation expresses the opinion of the presenter and may not be understood or quoted as being made on behalf of or reflecting the position of the European Medicines Agency or one of its committees or working parties.

Outline

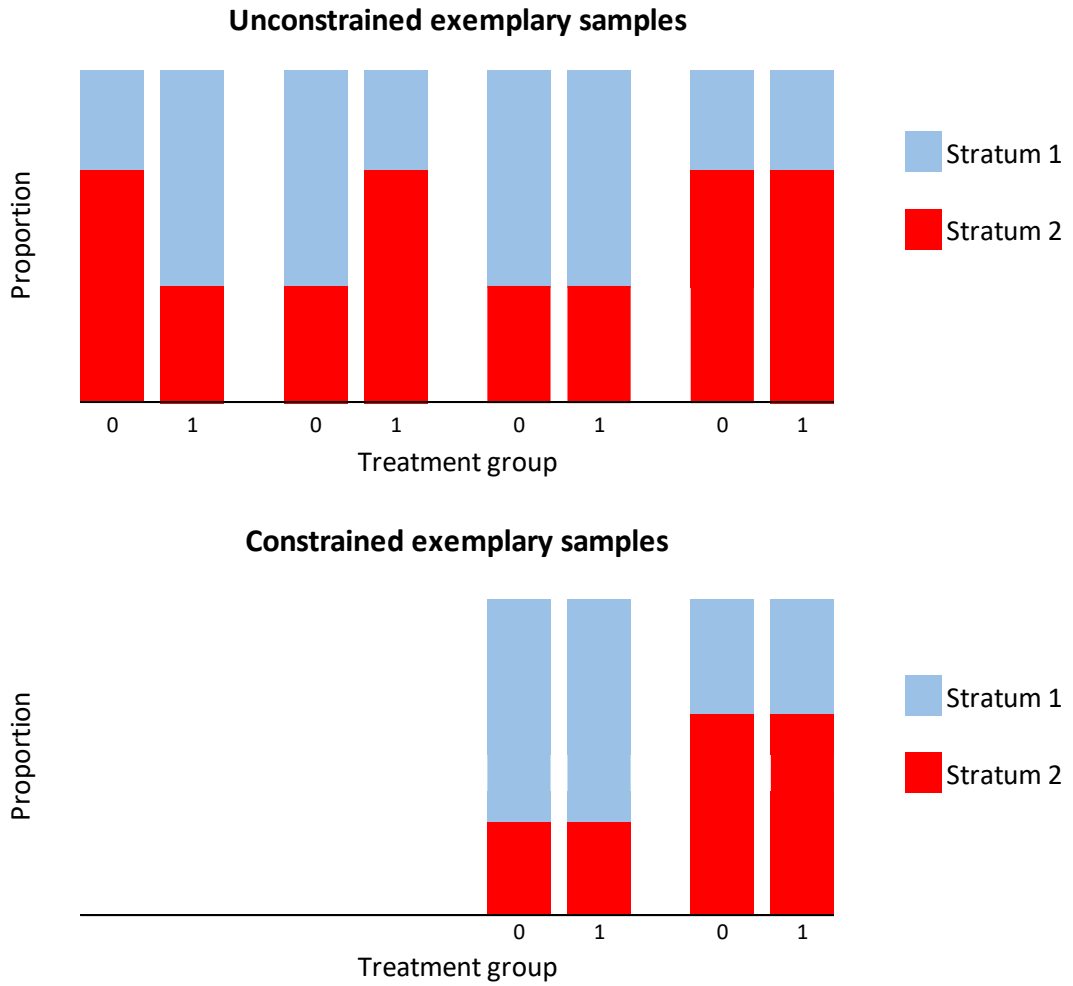
- Role of baseline covariates in randomized controlled trials
- Current guidance on covariate adjustment
- Marginal versus conditional estimands
- Number of covariates and finite sample properties
- Prespecification
- Sample size calculations
- Complex designs

Guidelines

- ICH E9 and addendum on estimands
- EMA Guideline on adjustment for baseline covariates in clinical trials
- FDA Adjusting for Covariates in Randomized Clinical Trials for Drugs and Biological Products

Role of baseline covariates – Design stage

- Constrained randomization (e.g. stratified permuted block)
- Balance covariate distribution between treatment groups
- Reduce variance of observed between group difference: Correlation between the per-group estimates, orthogonal design matrix [Kahan & Morris 2012, Senn et al. 2026]
- Need to account for in analysis to fully harvest efficiency gain)
- Credibility: Imbalance may raise concerns of confounding conditional on given data set
- Constrained randomization biggest impact with small sample sizes



Role of baseline covariates – Analysis stage

- Explain variability, improve precision
- Adjust for imbalance at analysis stage: Minor role, needs model assumptions
- In conditional estimands: Interpret effect as more patient specific, extrapolation to populations with different covariate distribution controversially discussed [Spieker 2022, Van Lancker et al 2024], needs model assumptions
- When stratified randomization uses categorized continuous variables, either version or both may be included as covariate. Continuous version often preferable [Senn et al. 2026].
- It is in general not of interest to analyse the association between covariates and outcome

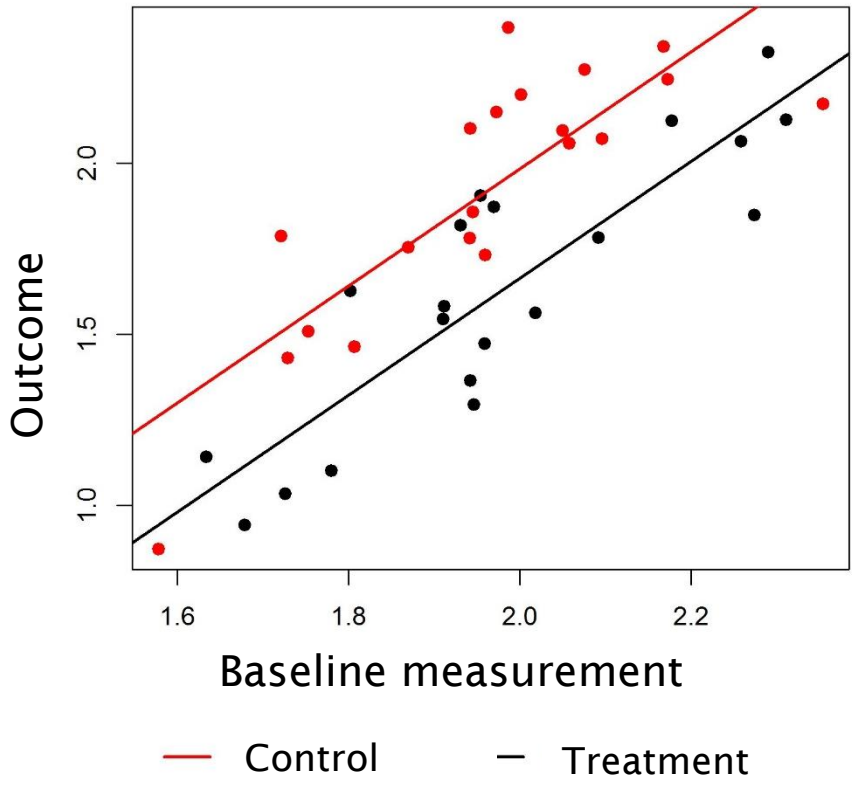


Illustration of ANCOVA explaining variance of outcome through baseline measurments

Traditional focus for covariate adjusted analyses

- Guidelines uniformly recommend utilizing covariates to gain precision
- Caveat in EMA guideline: *„smaller P-values may not be sufficient to produce convincing evidence of a clinically useful effect: the size of the treatment effect and its consistency across levels of covariates will always be important considerations."*
- **Continuous endpoints:**
  - ANCOVA favoured approach
  - No treatment-by-covariate interactions
  - Robust with respect to assumptions (may need sandwich variance for other than 1:1 randomization)
  - Asymptotic efficiency gain over unadjusted analysis
- **Binary and survival endpoints:**
  - Stratified analysis to account for stratified randomization
  - Additional covariates possible in logistic or Cox models
  - Impact on efficiency less obvious
- Common understanding that stratified randomization should be reflected in the analysis, otherwise variability overestimated

„New“ methods focussing on marginal estimands

- General covariate augmentation
- Covariates reduce variability of estimating equation
- Impact on efficiency more obvious than in GLMs
- Asymptotically, efficiency is improved [Bannick et al. 2025]
- Recent proposals to include more complex covariate information
  - ANHECOVA: ANCOVA with treatment by covariate interaction [e.g. Ye et al. 2023]
  - PROCOVA, use supercovariate defined with external data [Schuler et al. 2022]
  - Machine learning to replace linear predictors in general covariate adjustment, need to account for overfitting (e.g. cross fitting) [e.g. Van Lancker et al. 2026]

Potential impact of „new“ methods of marginal adjustment

- Continuous endpoints
  - Typically larger gain from covariate adjustement compared to binary/time-to-event (Shao 2026: median variance reduction of 13.3% for continuous outcomes and 4.6% for binary outcomes in 50 trials)
  - Potential for further improvement by complex models
  - Caveat: Trade-off between sample size and efficiency gain. Small samples may not support estimation of required model parameters.
- Binary endpoints
  - Focus on marginal estimation may facilitate interpretation
  - Compared to logistic regression: May choose summary measure independent of model assumptions
  - Compared to Cochrane-Mantel-Haenszel approach: Use additional/continuous covariates
- Survival
  - Impact less obvious
  - Marginal adjustment for hazard ratio still needs PH assumption for interpretation
  - Methods matching other summary measures (RMST, x-year survival) may be less powerful

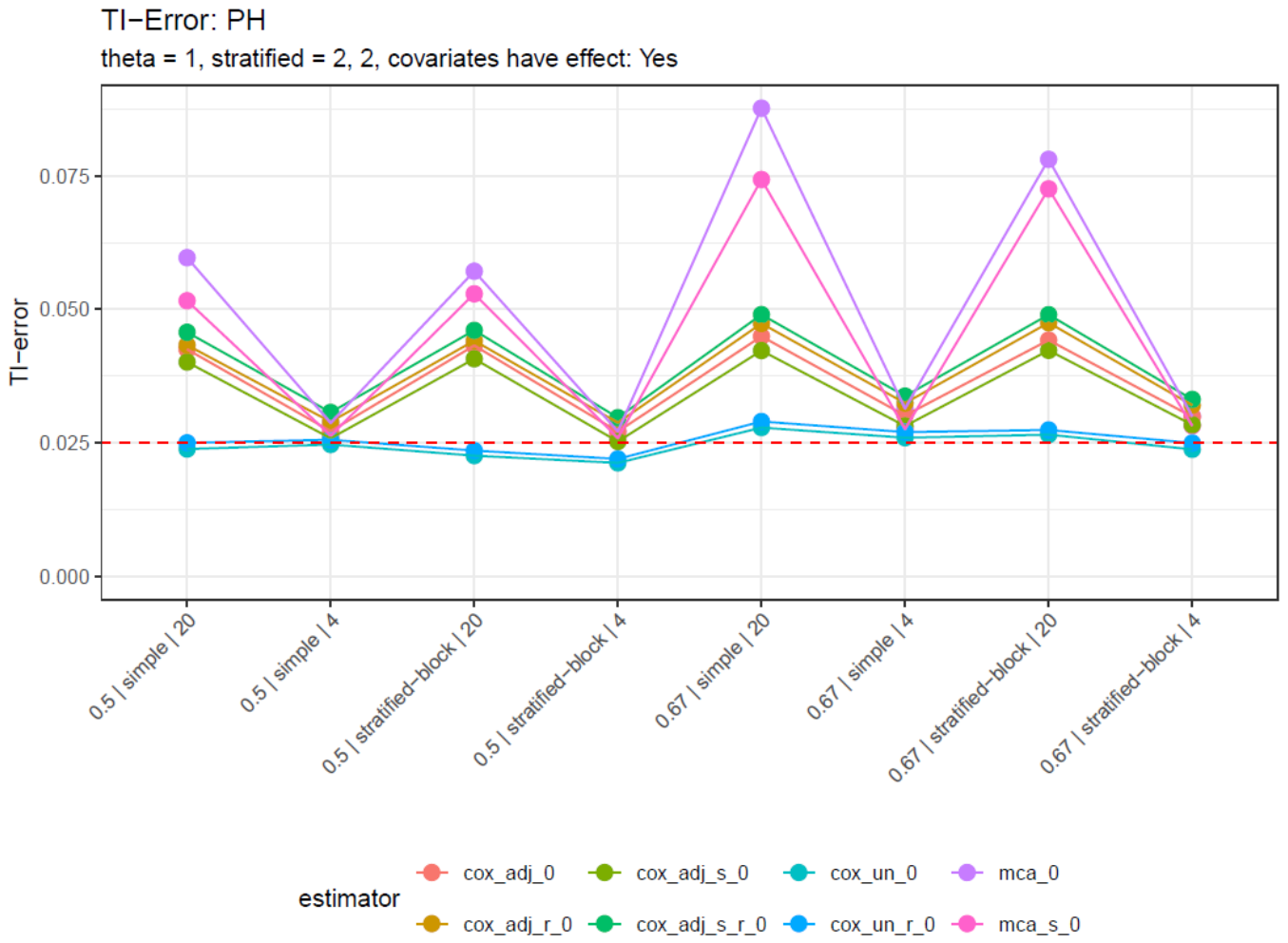
Marginal versus conditional estimands in guidelines

- EMA guideline
  - Focused on traditional models
  - Implies conditional estimands in case of non-linear models
  - Does not rule out marginal estimands
- FDA guideline
  - Both approaches in principle acceptable
  - Detailed guidence on marginal estimands and estimation methods
  - Allows for unadjusted analysis
- Align analysis with estimand

(See, e.g., Daniel et al. 2021 for discussion of marginal vs. conditional estimands and unadjusted vs. adjusted estimation.)

Number of covariates

- Traditional guidance: Few covariates
- ICH E9 and FDA guideline mention unadjusted analysis as viable option
- Methods based solely on large sample asymptotics do not account for uncertainty introduced by additional model parameters
  - Type I error inflation
- With proper finite sample corrections (akin to degrees of freedom in linear models), too many covariates may reduce power
- Few covariates may remain advisable approach



Simulated type I error rates in RCT with time-to-event endpoint and 66 events.
x-axis: Allocation ratio | randomization | number of covariates
Methods: Unadjusted Cox models (cox_un), conditionally adjusted Cox (cox_adj), marginally adjusted Cox (mca) as in Ye et al. 2024. r...robust sandwich variance, s...stratified.

Importance of prespecification

- New methods offer increased complexity
  - Potentially larger covariate space
  - Functional shapes other than linear
  - Arbitrarily composed supercovariates
- More technical parameters to set
  - Decisions when implementing marginal adjusted logrank test proposed by Ye et al. 2024:
    - Use expected or observed group allocation ratio in the variance estimator?
    - Fit internal linear model once, or update while iteratively fitting the hazard ratio?
  - With cross-fitting, which split to choose, how to guarantee reproducible results?
  - Choice between different variance estimators
- Need to prespecify models exactly
- Reviewers need to be aware of modelling options to be able to assess whether prespecification is complete

Sample size calculation

- Should the potential gain in efficiency be accounted for in sample size planning?
  - May lead to smaller (faster, more cost effective) trials
- Traditionally, covariate information is not used in sample size calculations
- Power may be typically larger than planned 80% or 90%
- FDA guideline „*sample size and power calculations can be based on adjusted or unadjusted methods*"
- But: Requires good understanding of impact of covariates on sample size/power
  - Straight forward for ANCOVA, less obvious for general covariate adjustment
  - Risk of too optimistic assumptions on effect of covariates
- Sample size should be sufficient to address other aspects than primary endpoint
  - Safety data base
  - Secondary endpoints
  - Rely on large sample asymptotics
- Conservative sample size planning remains advisable

Complex study designs

- Many trials include interim analyses or sample size reassessment
- Stopping boundaries are traditionally calculated based on independent increment property of group sequential statistics, which implies a certain covariance structure
- This property may not hold with covariate adjustment [Van Lancker et al. 2025, Bannick et al. 2026]
- Methods to adjust decision rules have been proposed [Van Lancker et al. 2025, Backenroth et al. 2026]
  - Based on estimated covariances
  - Small sample properties may warrant further investigations
- Similar concerns may arise in the context of other complex designs, e.g. conditional power calculations for sample size reassessment

Simulation Example: 1:1 simple randomization, survival endpoint, max. 380 events, 3 interim looks, Pocock-type alpha spending, four covariates, Ye‘s marginally adjusted logrank test

Type I error inflated to 2.85% (simulation SE 0.05%)

Covariance matrix of score statistics:

	S1	S2	S3
S1	**29.1**	**27.2**	**25.8**
S2	27.2	55.5	53.0
S3	25.8	53.0	81.3

Apprixomate variance of unadjusted score statistics:

31.7 63.3 95.0

Software

- For new methods, limited availability of software implementations
  - Development of R packages RobinCar, RobinCar2 [Bannick et al. 2026], beeca acknowledged
- Ideally validate software with independent implementation

Summary of regulatory challenges

- For each particular study, need to agree on estimand
  - Marginal vs. Conditional
  - Summary measure
- Concerns regarding finite sample properties:
  - Type I error control?
  - Is a complex model indeed more efficient than simple or traditional model?
  - In supporting simulation studies, need to ensure that the simulation settings cover sufficient deviations from assumptions.
- Proporties in combination with specific study design
  - Randomization procedure
  - Interim analysis
  - Sample size reassessment
- Need to be aware of analysis model options to assess prespecification
- Valid software implementation

Literature

- Backenroth, D., Roels, S., Dibaj, S., Ye, T., Öhrn, F., & Van Lancker, K. (2026). Practical considerations when using the covariate-adjusted log-rank test for the analysis of time-to-event endpoints in oncology trials. Biometrics, 82(2), ujag109.
- Bannick, M. S., Shao, J., Liu, J., Du, Y., Yi, Y., & Ye, T. (2025). A general form of covariate adjustment in clinical trials under covariate-adaptive randomization. Biometrika, 112(3), asaf029.
- Bannick, M. S., Yi, Y., & Ye, T. (2026). Covariate adjustment in randomized clinical trials: From general theory to practical insights. Clinical Trials, 23(3), 272-281.
- Bannick, M., Bian, Y., Chen, G., Li, L., Qian, Y., Bové, D. S., ... & Yi, Y. (2026). The RobinCar Family: R Tools for Robust Covariate Adjustment in Randomized Clinical Trials. arXiv preprint arXiv:2601.14498.
- Daniel, R., Zhang, J., & Farewell, D. (2021). Making apples from oranges: Comparing noncollapsible effect estimators and their standard errors after adjustment for different covariate sets. Biometrical Journal, 63(3), 528-557.
- Kahan, B. C., & Morris, T. P. (2012). Improper analysis of trials randomised using stratified blocks or minimisation. Statistics in medicine, 31(4), 328-340.
- Schuler, A., Walsh, D., Hall, D., Walsh, J., Fisher, C., Critical Path for Alzheimer's Disease, ... & Alzheimer's Disease Cooperative Study. (2022). Increasing the efficiency of randomized trial estimates via linear adjustment for a prognostic score. The International Journal of Biostatistics, 18(2), 329-356.
- Senn, S., König, F., & Posch, M. (2026). Stratification in small randomised clinical trials and analysis of covariance: Some simple theory and recommendations. Statistical Methods in Medical Research, 09622802261454833.
- Shao, Y., Lyu, L., Yu, M., & Wang, B. (2026). How should covariates be handled in randomized trials? Empirical evidence from 50 trials and recommendations for practice. Journal of Clinical Epidemiology, 112374.
- Spieker, A. J. (2022). Comments on the debate between marginal and conditional estimands. Statistics in medicine, 41(28), 5589-5591.
- Van Lancker, K., Bretz, F., & Dukes, O. (2024). Covariate adjustment in randomized controlled trials: General concepts and practical considerations. Clinical Trials, 21(4), 399-411.
- Van Lancker, K., Díaz, I., & Vansteelandt, S. (2026). Automated, efficient and model-free inference for randomized clinical trials via data-driven covariate adjustment. The American Statistician, (just-accepted), 1-26.
- Ye T, Shao J, Yi Y, et al. Toward better practice of cov-ariate adjustment in analyzing randomized clinical trials. J Am Stat Assoc 2023; 118(544): 2370–2382.
- Ye, T., Shao, J., & Yi, Y. (2024). Covariate-adjusted log-rank test: guaranteed efficiency gain and universal applicability. Biometrika, 111(2), 691-705.