Adjusting for Covariates in Randomized Clinical Trials

Dalong Patrick Huang

FDA/CDER/OB/DBVI

FDA U.S. FOOD & DRUG ADMINISTRATION

Disclaimer

This presentation reflects the views of the author and should not be construed to represent views or policies of U.S. Food and Drug Administration.

FDA Guidance Documents on Covariate Adjustment

- ICH E9 Statistical Principles for Clinical Trials (1998):

  "Pretrial deliberations should identify those covariates and factors expected to have an important influence on the primary variable(s), and should consider how to account for these in the analysis to improve precision..."

- COVID-19: Developing Drugs and Biological Products for Treatment or Prevention draft guidance (2020)

  "To improve the precision of treatment effect estimation and inference, sponsors should consider adjusting for prespecified prognostic baseline covariates (e.g., age, baseline severity, comorbidities, baseline medications and COVID-19 vaccination status) in the primary efficacy analysis and should propose methods of covariate adjustment."

- **Adjusting for Covariates in Randomized Clinical Trials for Drugs and Biological Products revised draft guidance (2021)**

  The main focus of the guidance is on the use of prognostic baseline factors to improve precision for estimating treatment effects.

Adjusting for Covariates, Revised Draft Guidance (2021)

- I. INTRODUCTION
- II. BACKGROUND
- **III. RECOMMENDATIONS FOR COVARIATE ADJUSTMENT IN RANDOMIZED TRIALS**
  - **A. General Considerations**
  - **B. Linear models**
  - **C. Nonlinear models**
- IV. REFERENCES

General Considerations

- Adjustment is recommended because it often improves power and precision, and unadjusted analysis remains acceptable as well.

- Prespecification: covariates used for adjustment and mathematical form of the model

- Adjusting for covariates that are prognostic for the outcome leads to the greatest efficiency gains
  - Scientific literature
  - Previous studies (e.g., a Phase 2 trial)
  - Properties of adjustment are best understood when the number of covariates is small relative to the sample size (Tsiatis et al. 2008)

- Stratified randomization
  - Analysis ignoring stratified randomization is likely to overestimate standard errors (SEs) and can be unduly conservative for inference
  - Recommend SE computation account for stratified randomization (Bugni et al. 2018; Ye et al. 2021)

Linear Model: Estimation

- Estimand is population average treatment effect (i.e., difference in expected outcomes between subjects assigned to treatment and control groups)

- Usual adjusted estimator is least squares fit of treatment coefficient in a regression of the outcome on an intercept, treatment, and baseline covariates

- Can provide valid estimation of the average treatment effect in a randomized trial even when the linear model is misspecified (Lin, 2013)

Linear Model: Robust Standard Error

- Nominal SEs reported by most packages for generalized linear models can be inaccurate if the model is incorrect

- Otherwise, can be corrected with robust SE
  - Huber-White “sandwich” SEs when model does not include treatment by covariate interaction (Rosenblum and van der Laan 2009, Lin 2013)
  - Other robust SEs for linear model with interactions (Ye et al. 2022)
  - Appropriate nonparametric bootstrap procedure (Efron and Tibshirani 1993)

Linear Model: Interactions

- The linear model may include treatment by covariate interaction terms

- However, when using this approach, the primary analysis should still be based on an estimate from the model of the average treatment effect (Tsiatis et al. 2008; Ye et al 2021)

- Per ICH E9, interaction effects may be important to assess in supportive analysis or exploratory analysis because differences in treatment effects across subgroups defined by baseline covariates could be relevant to prescribers, patients, and other stakeholders and imply that the average treatment effect gives an incomplete summary of efficacy

Linear Model: Example

- Estimand: Difference in mean FEV1 at 12 weeks between drug and placebo in patients with moderate-to-severe asthma regardless of adherence to treatment or use of ancillary medications

- Main analysis: ANCOVA of FEV1 at 12 weeks in all randomized patients, adjusting for baseline FEV1, age, and sex, with Huber-White sandwich standard errors

Nonlinear Model: Collapsibility

- With binary, ordinal, or time-to-event outcomes certain population-level summaries can be non-collapsible even in randomized trials

Table 1: Non-collapsibility of the odds ratio in a hypothetical target population

	Percentage of target population	Success rate		Odds ratio
		New drug	Placebo	
Males	50%	80.0%	33.3%	8.0
Females	50%	25.0%	4.0%	8.0
Combined	100%	52.5%	18.7%	4.8

- As part of the prespecification of the estimand of intertest, sponsors should specify whether treatment effect of interest is the unconditional (e.g., 4.8 in the table) or conditional treatment effect (e.g., 8.0 in the table)

Nonlinear Model: Conditional Treatment Effects

- Nonlinear models such as logistic regression or proportional hazards regression for conditional treatment effects are commonly used in clinical trials

- Advantages:
  - Can provide more personalized information than unconditional treatment effects if assumption holds (and not otherwise)
  - Available in standard statistical software packages
- Disadvantage:
  - When estimating a conditional treatment effect through nonlinear regression, the model will generally not be exactly correct, and results can be difficult to interpret if the model is misspecified and treatment effects substantially differ across subgroups.

- Sponsors should discuss with the relevant review divisions specific proposals in a protocol or statistical analysis plan containing nonlinear regression to estimate conditional treatment effects for the primary analysis

Nonlinear Model: Unconditional Treatment Effects

- Sponsors can perform covariate adjusted estimation and inference for an unconditional treatment effect in the primary analysis
- The estimand will be the same as in an unadjusted analysis
- The method used should provide valid inference under approximately the same minimal statistical assumptions that would be needed for unadjusted estimation
- Statistically reliable methods
  - Binary outcomes (e.g., Ge et al. 2011)
  - Ordinal outcomes (e.g., Díaz et al. 2016)
  - Time-to-event outcomes (e.g., Tangen and Koch 1999; Lu and Tsiatis 2008)
- SEs or confidence intervals can be formed from the nonparametric bootstrap or formulas justified in the statistical literature (Colantunoni and Rosenblum 2015)

Example of "Standardized", "Plug-in", or G-computation" Estimator for Unconditional Effect with Binary Outcomes (Ge et al. 2011)

For Binary outcome Y, treatment A (1=treatment, 0=control), covariate B:

- Fit logistic regression model for

$P(Y = 1|A, B) = \text{logit}^{-1}(\beta_0 + \beta_1 A + \beta_2 B)$.

- Compute standardized estimators for treatment specific means μ_0, μ_1:

$$\hat{\mu}_0 = \frac{1}{n} \sum_{i=1}^{n} \text{logit}^{-1}(\hat{\beta}_0 + \hat{\beta}_2 B_i)$$
$$\hat{\mu}_1 = \frac{1}{n} \sum_{i=1}^{n} \text{logit}^{-1}(\hat{\beta}_0 + \hat{\beta}_1 + \hat{\beta}_2 B_i)$$

- Estimator is contrast of interest between μ_1, μ_0, e.g., risk difference $\hat{\mu}_1 - \hat{\mu}_0$.

Same holds for other (unconditional) estimands, e.g., relative risk

Nonlinear Model: Unconditional Effect Example

- Estimand: Difference in probability of 28-day survival between drug and placebo in severe-to-critical hospitalized COVID-19 patients regardless of adherence to treatment or use of ancillary medications

- Main Analysis: A logistic model in all randomized patients adjusting for age, baseline severity, and COVID-19 vaccination status, with a standardized (plug-in) estimator of the risk difference and SE (Ge et al. 2011) and a Wald test

Acknowledgments

- Dr. Dan Rubin
- Professor Michael Rosenblum
- CDER/OB Covariate Adjustment Working Group

Reference

- ICH E9 Statistical Principles for Clinical Trials, **Step 4** version, 5 February 1998 https://database.ich.org/sites/default/files/E9_Guideline.pdf
- COVID-19: Developing Drugs and Biological Products for Treatment or Prevention draft guidance, 2020 https://www.fda.gov/media/137926/download
- Adjusting for Covariates in Randomized Clinical Trials for Drugs and Biological Products revised draft guidance, 2021 https://www.fda.gov/media/148910/download
- Bugni F, Canay IA, and AM Shaikh, 2018, Inference Under Covariate-Adaptive Randomization, Journal of the American Statistical Association, 113(524):1784-1796.
- Colantuoni E, and M Rosenblum, 2015, Leveraging prognostic baseline variables to gain precision in randomized trials, Statistics in Medicine, DOI: 10.1002/sim.6507.
- Díaz I, Colantuoni E, and M Rosenblum, 2016, Enhanced precision in the analysis of randomized trials with ordinal outcomes, Biometrics, 72(2):422-431.
- Efron B and RJ Tibshirani, 1993, An Introduction to the Bootstrap, Boca Raton (FL): Chapman & Hall.
- Freedman DA, 2008, Randomization Does Not Justify Logistic Regression, Statistical Science, 23(2):237-249.
- Steingrimsson, Hanley, and Rosenblum, 2017, Improving precision by adjusting for prognostic baseline variables in randomized trials with binary outcomes, without regression model assumptions, Contemporary Clinical Trials, 54:18-24.

Reference

- Ge M, Durham LK, Meyer RD, Xie W, and Thomas N, 2011, Covariate-Adjusted Difference in Proportions from Clinical Trials Using Logistic Regression and Weighted Risk Differences, Drug Information Journal, 45(4):481-493.

- Lin W, 2013, Agnostic Notes on Regression Adjustments to Experimental Data: Reexamining Freedman's Critique, Annals of Applied Statistics, 7(1):295-318.

- Lu X and AA Tsiatis, 2008, Improving the Efficiency of the Log-Rank Test Using Auxiliary Covariates, Biometrika, 95(3):679-694.
- Rosenblum M and MJ van der Laan, 2009, Using Regression Models to Analyze Randomized Trials: Asymptotically Valid Hypothesis Tests Despite Incorrectly Specified Models, Biometrics, 65(3):937-945.
- Rosenblum M and MJ van der Laan, 2010, Simple, efficient estimators of treatment effects in randomized trials using generalized linear models to leverage baseline variables, International Journal of Biostatisitcs, 6, Article 13.

- Tangen CM and GG Koch, 1999, Non-Parametric Analysis of Covariance for Hypothesis Testing with Logrank and Wilcoxon Scores and Survival-Rate Estimation in a Randomized Clinical Trial, Journal of Biopharmaceutical Statistics, 9(2):307-338.

- Tsiatis AA, Davidian M, Zhang M, and X Lu, 2008, Covariate Adjustment for Two-Sample Treatment Comparisons in Randomized Trials: A Principled Yet Flexible Approach, Statistics in Medicine, 27(23):4658-4677.

- Ye T, Yi Y, and Shao J. Inference on the average treatment effect under minimization and other covariate-adaptive randomization methods. Biometrika, 2021. https://doi.org/10.1093/biomet/asab015
- Ye T, Shao J, Yi Y, and Zhao Q. Toward better practice of covariate adjustment in analyzing randomized clinical trials. Journal of the American Statistical Association, 2022. https://doi.org/10.1080/01621459.2022.2049278

Backup Slides

Simulation Study Result: difference in restricted mean survival times (RMST) 14 days after hospitalization

n	Estimator	Power	MSE	RE
100	Unadjusted	0.09	53.7	1.00
100	Adjusted	0.15	51.0	0.95
200	Unadjusted	0.33	62.7	1.00
200	Adjusted	0.40	56.4	0.90
500	Unadjusted	0.74	72.9	1.00
500	Adjusted	0.82	62.2	0.85
1000	Unadjusted	0.96	76.5	1.00
1000	Adjusted	0.98	63.5	0.83

n=sample size; RE=relative efficiency (ratio of adjusted vs. unadj. MSE).

Benkeser et al. (2020)

Nonlinear Model: Precision and Efficiency of the Unadjusted, Standardized, and Logistic Coefficient Estimators

	Estimator	Average value of estimator	Empirical standard error	Relative efficiency	Reduction in sample size
Setting 1	Unadjusted	0.13	4.5×10^{-2}	1	0
	Standardized	0.13	3.8×10^{-2}	1.41	29%
	Logistic coefficient	0.76	0.23	1.31	24%
Setting 2	Unadjusted	0.13	4.5×10^{-2}	1	0
	Standardized	0.13	4.5×10^{-2}	0.99	−1%
	Logistic coefficient	0.53	0.19	0.94	−7%

Setting 1: Baseline variables prognostic for the outcome; Setting 2: Baseline variables independent of the outcome; For both the unadjusted and standardized estimator, the true unconditional treatment effect is 0.13 in both settings. In setting 2, the true conditional treatment effect on the log odds scale is 0.52. As the logistic regression model is not necessarily correct in setting 1, it is unclear if the true conditional effect is interpretable as a single number.

Steingrimsson, Hanley, and Rosenblum 2017

Nonlinear Model: Properties of the Unadjusted, Standardized, and Logistic Coefficient Estimators

Estimator	Effect it estimates	Requires regression model assumptions?	Adjusts for baseline variables?
Unadjusted	Unconditional effect	No	No
Standardized	Unconditional effect	No	Yes
Logistic coefficient	Conditional effect	Yes	Yes

Steingrimsson, Hanley, and Rosenblum 2017

Nonlinear Model: Example of Misspecified Logistic Model

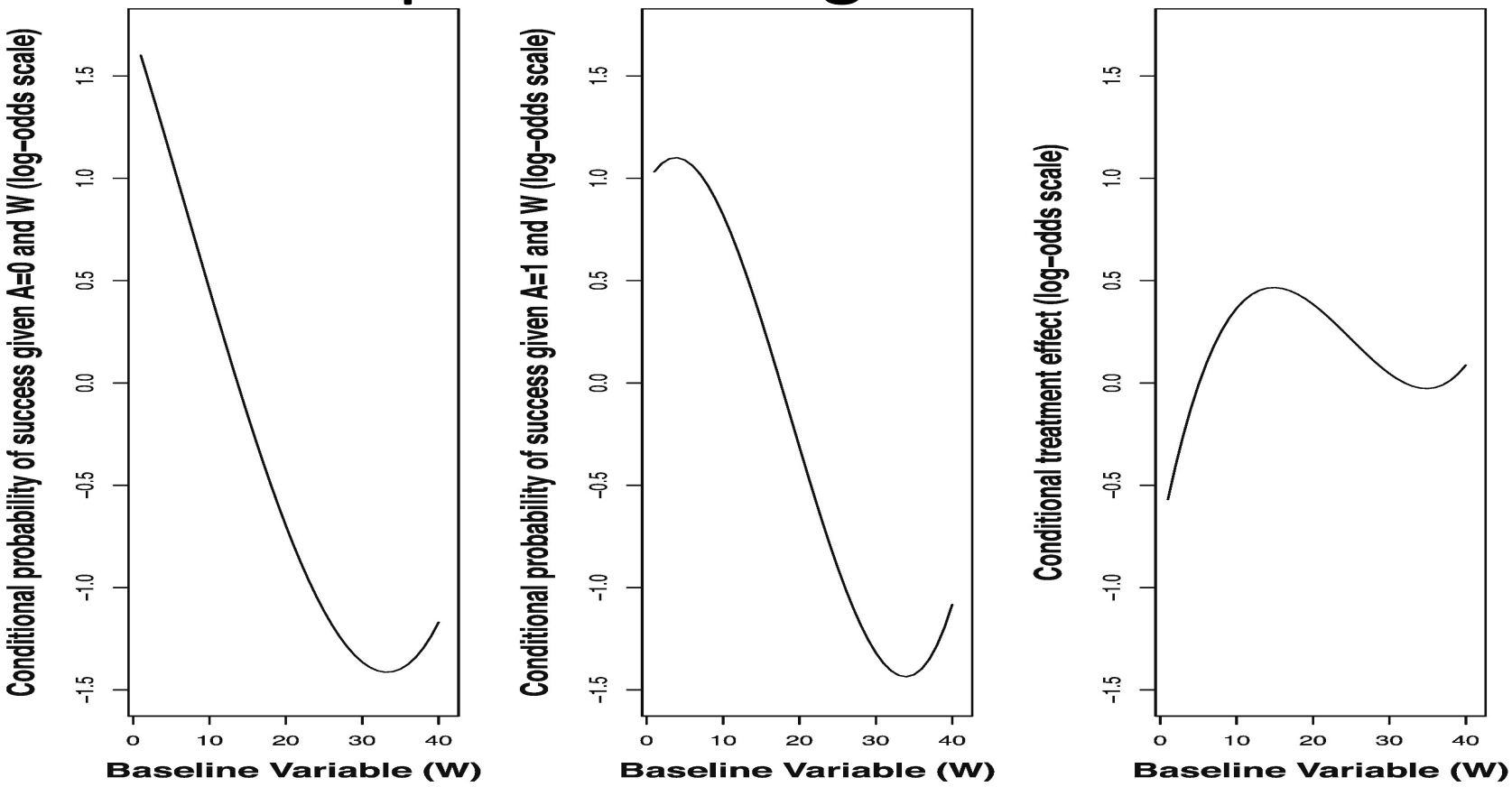
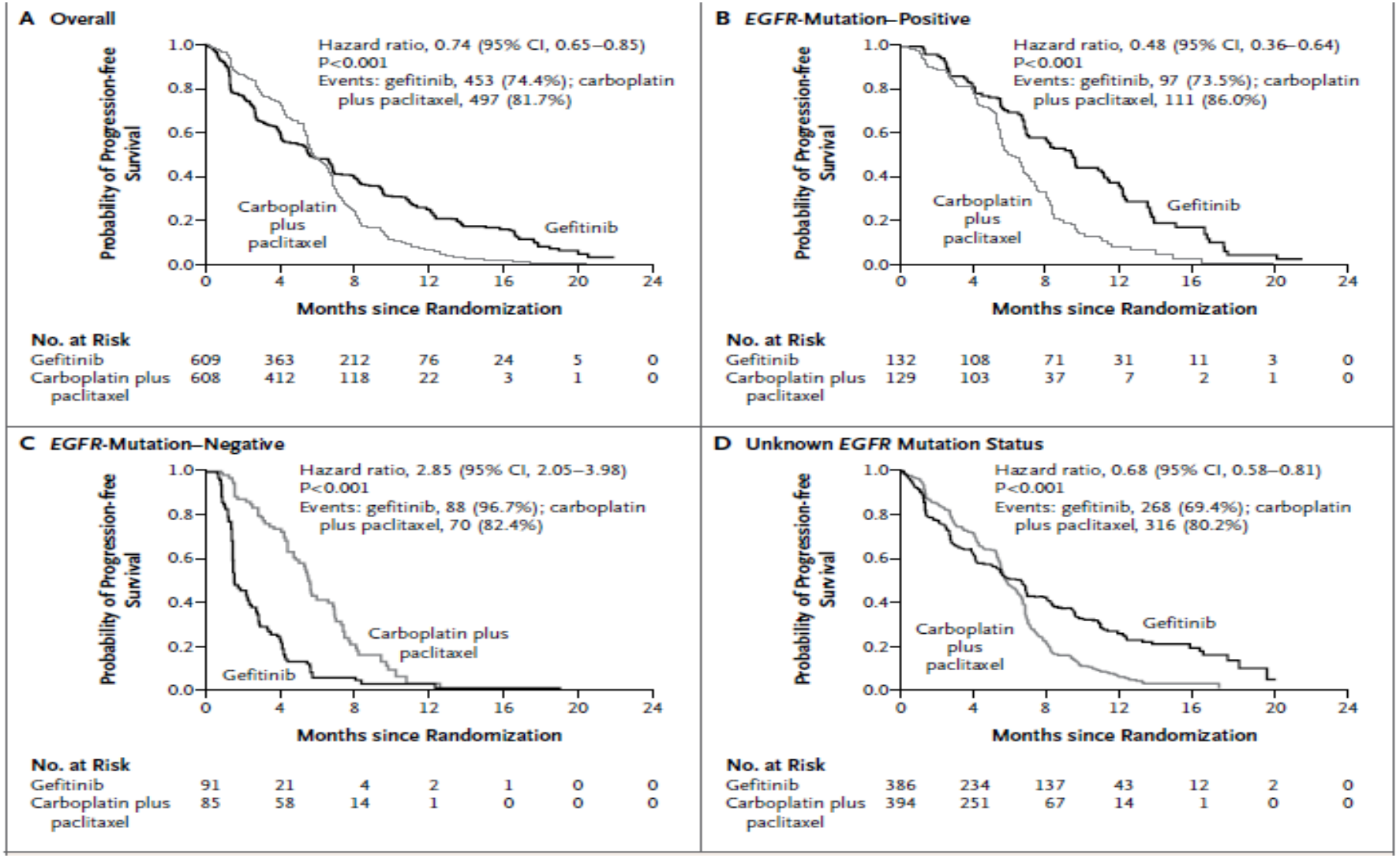


Fig. 1. An example of a conditional effect that depends on the value of the baseline variables and cannot be represented using a single number.

Steingrimsson, Hanley, and Rosenblum 2017

Non-proportional Hazard Ratio Example



Mok et al. 2009 NEMJ


FDA
U.S. FOOD & DRUG
ADMINISTRATION