Adjusting for Covariates in Randomized Clinical Trials for Drugs and Biological Products

Guidance for Industry

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

Comments and suggestions regarding this draft document should be submitted within 60 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Submit electronic comments to https://www.regulations.gov. Submit written comments to the Dockets Management Staff (HFA-305), Food and Drug Administration, 5630 Fishers Lane, Rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.

For questions regarding this draft document, contact (CDER) Scott N. Goldie at 301-796-2055, or (CBER) Office of Communication, Outreach and Development, 800-835-4709 or 240-402-8010.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

May 2021
Biostatistics

Revision 1

Adjusting for Covariates in Randomized Clinical Trials for Drugs and Biological Products

Guidance for Industry

Additional copies are available from:
Office of Communications, Division of Drug Information
Center for Drug Evaluation and Research
Food and Drug Administration
10001 New Hampshire Ave., Hillandale Bldg., 4th Floor
Silver Spring, MD 20993-0002
Phone: 855-543-3784 or 301-796-3400; Fax: 301-431-6353
Email: druginfo@fda.hhs.gov
https://www.fda.gov/drugs/guidance-compliance-regulatory-information/guidances-drugs
and/or
Office of Communication, Outreach and Development
Center for Biologics Evaluation and Research
Food and Drug Administration
10903 New Hampshire Ave., Bldg. 71, Room 3128
Silver Spring, MD 20993-0002
Phone: 800-835-4709 or 240-402-8010
Email: ocod@fda.hhs.gov
https://www.fda.gov/vaccines-blood-biologics/guidance-compliance-regulatory-information-biologics/biologics-guidances

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

May 2021
Biostatistics

Revision 1

TABLE OF CONTENTS

Adjusting for Covariates in Randomized Clinical Trials for Drugs and Biologics
Guidance for Industry[1]

This draft guidance, when finalized, will represent the current thinking of the Food and Drug Administration (FDA or Agency) on this topic. It does not establish any rights for any person and is not binding on FDA or the public. You can use an alternative approach if it satisfies the requirements of the applicable statutes and regulations. To discuss an alternative approach, contact the FDA staff responsible for this guidance as listed on the title page.

I. INTRODUCTION

This guidance represents FDA's current thinking on adjusting for covariates in the statistical analysis of randomized clinical trials in drug[2] development programs. This guidance provides recommendations for the use of covariates in the analysis of randomized, parallel group clinical trials that are applicable to both superiority trials and noninferiority trials. The main focus of the guidance is on the use of prognostic baseline factors[3] to improve precision for estimating treatment effects rather than the use of predictive biomarkers to identify groups more likely to benefit from treatment. This guidance does not address use of covariates to control for confounding variables in non-randomized trials or the use of covariate adjustment for analyzing longitudinal repeated measures data.

This guidance revises the draft guidance for industry *Adjusting for Covariates in Randomized Clinical Trials for Drugs and Biologics with Continuous Outcomes* issued in April 2019. This revision provides more detailed recommendations for the use of linear models for covariate adjustment and also includes recommendations for covariate adjustment using nonlinear models.

The contents of this document do not have the force and effect of law and are not meant to bind the public in any way, unless specifically incorporated into a contract. This document is intended only to provide clarity to the public regarding existing requirements under the law. FDA guidance documents, including this guidance, should be viewed only as recommendations, unless specific regulatory or statutory requirements are cited. The use of the word *should* in Agency guidances means that something is suggested or recommended, but not required.

[1] This guidance has been prepared by the Office of Biostatistics in the Center for Drug Evaluation and Research in cooperation with the Center for Biologics Evaluation and Research at the Food and Drug Administration.

[2] The term *drug* used in this guidance refers to both human drugs and biological products.

[3] The term *prognostic baseline factors* used in this guidance refers to baseline covariates that are likely to be associated with the primary endpoint.

II. BACKGROUND

Baseline covariates in this guidance refer to demographic factors, disease characteristics, or other information collected from participants before the time of randomization. Covariate adjustment refers to the use of baseline covariate measurements for estimating and testing treatment effects between randomized groups.

The target population for a new drug usually includes patients with diverse prognostic baseline factors. A randomized controlled trial can be used to estimate treatment effects even if the primary analysis does not consider these baseline covariates (through what is termed an unadjusted analysis) because measured and unmeasured covariates will on average be balanced between treatment groups. However, incorporating prognostic baseline factors in the primary statistical analysis of clinical trial data can result in a more efficient use of data to demonstrate and quantify the effects of treatment with minimal impact on bias or the Type I error rate.

The ICH guidance for industry *E9 Statistical Principles for Clinical Trials (September 1998)*[4] addresses these issues briefly. The ICH E9 guidance encourages the identification of "covariates and factors expected to have an important influence on the primary variable(s)." The ICH E9 guidance strongly advises prespecification of "the principal features of the eventual statistical analysis," including "how to account for [covariates] in the analysis to improve precision and to compensate for any lack of balance between treatment groups." The ICH E9 guidance also cautions against adjusting for "covariates measured after randomization because they could be affected by the treatments."

This guidance provides general considerations and additional recommendations for covariate adjustment using linear and nonlinear models. In linear models, adjustment for baseline variables often leads to improved precision by reducing residual variance. When adjusting for covariates based on fitting nonlinear regression models, such as logistic regression models in studies with binary outcomes, there are additional considerations that arise because inclusion of baseline covariates in a regression model can change the treatment effect that is being estimated. As explained below, after suitably addressing the treatment effect definition, covariate adjustment using linear or nonlinear models can be used to increase precision.

III. RECOMMENDATIONS FOR COVARIATE ADJUSTMENT IN RANDOMIZED TRIALS

A. General Considerations

- Sponsors can adjust for baseline covariates in the analyses of efficacy endpoints in randomized clinical trials.

[4] We update guidances periodically. To make sure you have the most recent version of a guidance, check the FDA guidance web page at https://www.fda.gov/regulatory-information/search-fda-guidance-documents.

- Although an unadjusted analysis is acceptable for the primary analysis, adjustment for baseline covariates will generally reduce the variability of estimation of treatment effects and thus lead to narrower confidence intervals and more powerful hypothesis testing.

- Sponsors should prospectively specify the covariates and the mathematical form of the covariate adjusted estimator in the statistical analysis plan before any unblinding of comparative data. FDA will generally give more weight in review to the prespecified primary analysis than to post-hoc analyses using different models or covariates.

- Covariate adjustment leads to efficiency gains when the covariates are prognostic for the outcome of interest in the trial. Therefore, the covariates FDA recommends for adjustment should be those that are anticipated to be most strongly associated with the outcome of interest. Covariate adjustment can still be performed with covariates that are not prognostic, but there may not be any gain in precision (or may be a loss in precision) compared with an unadjusted analysis.

- Covariate adjustment is generally robust to the handling of subjects with missing baseline covariates. Missing baseline covariate values can be singly or multiply imputed, or missingness indicators (Groenwold et al. 2012) can be added to the model used for covariate adjustment. Sponsors should not perform imputation separately for different treatment groups, and sponsors should ensure that imputed baseline values are not dependent on any post-baseline variables, including the outcome.

- For adjusted estimation based on linear models or generalized linear models, FDA recommends that sponsors estimate standard errors using the Huber-White robust "sandwich" estimator (Rosenblum and van der Laan 2009) or the nonparametric bootstrap method (Efron and Tibshirani 1993) rather than using nominal standard errors, which can be inaccurate if the model is incorrectly specified and which are often the default method for estimating standard errors in most statistical software packages.

- The statistical properties of covariate adjustment are best understood when the number of covariates adjusted for in the study is small relative to the sample size (Tsiatis et al. 2008). If the number of covariates is large relative to the sample size sponsors should provide a justification for their proposal.

- Randomization is often stratified by baseline covariates. In this case, FDA recommends that the standard error computation account for the stratified randomization (Bugni et al. 2018) with or without strata variables in an adjustment model. Otherwise, the standard error is likely to be overestimated and interval estimation and hypothesis testing can become unduly conservative.

- Covariate adjustment is acceptable even if baseline covariates are strongly associated with each other (e.g., body weight and body mass index). However, adjusting for less redundant variables generally provides greater efficiency gains.

- Clinical trials often record a baseline measurement of a defined characteristic and record a later measurement of the characteristic to be used as an outcome. When using this approach, adjusting for the baseline value rather than (or in addition to) defining the primary endpoint as a change from baseline is generally acceptable.

B. Linear models

- Covariate adjustment through a linear model is an acceptable method for analyzing data from a randomized clinical trial. Generally, the outcome is regressed on a treatment assignment indicator and baseline covariates using ordinary least squares, and the resulting estimated regression coefficient for the treatment indicator is the estimate of the treatment effect.

- Covariate adjustment through a linear model generally provides reliable estimation and inference for the average treatment effect, which is the difference in expected outcomes between subjects assigned to treatment and control groups. The average treatment effect is an example of an unconditional treatment effect, which quantifies the effect at the population level of moving a target population from untreated to treated. Covariate adjustment through a linear model is a valid method for estimating and performing inference for the average treatment effect even when the linear regression model does not fully capture the relationships between the outcome, treatment, and covariates (Lin 2013). However, the power of hypothesis tests and precision of estimates generally improves if the model more closely approximates the true relationships among the outcome, treatment, and covariates.

- Covariate adjustment through a linear model (without treatment by covariate interactions) also estimates a conditional treatment effect, which is a treatment effect assumed to be approximately constant across subgroups defined by baseline covariates in the model. The distinction between an average treatment effect and conditional treatment effect is often overlooked because they happen to coincide in linear models. These two types of treatment effects are discussed in more detail in section III.C.

- 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. As noted in the ICH E9 guidance, 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.

C. Nonlinear models

- Covariate adjustment with nonlinear models is often used in the analysis of clinical trial data when the primary outcome of interest is not measured on a continuous scale or is right censored (e.g., binary outcome, ordinal outcome, count outcome, or time-to-event outcome). Adjustment using nonlinear models is a potentially acceptable method for analyzing these data from a clinical trial. However, there are additional issues described below that should be considered before using nonlinear models.

- In general, treatment effects may differ from subgroup to subgroup. However, with some parameters such as odds ratios, even when all subgroup treatment effects are identical this subgroup-specific conditional treatment effect can differ from the unconditional treatment effect (i.e., the effect at the population level from moving the target population from untreated to treated) (Gail et al. 1984). This is termed non-collapsibility (Agresti 2002), which is distinct from confounding and can occur despite randomization and large sample sizes. An example of non-collapsibility of the odds ratio for a hypothetical clinical trial is illustrated in Table 1 below. The unconditional odds ratio in the hypothetical target population is 4.8, which is lower than the conditional odds ratio of 8.0 in each of the male and female subgroups. In trials with time-to-event outcomes, the hazard ratio is also generally non-collapsible. Unlike the odds ratio or hazard ratio, the risk difference and relative risk are collapsible.

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

- Cochran-Mantel-Haenszel methods (Mantel and Haenszel 1959) are acceptable for the analysis of clinical trial data and attempt to estimate a conditional treatment effect, which is assumed to be constant across subgroups defined by a covariate taking a discrete number of levels (e.g., the value 8.0 in Table 1).

- Fitting a nonlinear regression of the outcome on treatment and baseline covariates similarly attempts to estimate a conditional treatment effect. Nonlinear models extend Cochran-Mantel-Haenszel methods by allowing adjustment for continuous covariates, such as age. In nonlinear regression models (without treatment by covariate interactions) the treatment effect is assumed to be approximately constant across subgroups defined by baseline covariates in the model and can provide more personalized information than the unconditional treatment effect.

- While the adjusted estimator of a conditional odds ratio generally has a larger standard error than an unadjusted estimator of the unconditional odds ratio, this is not necessarily a disadvantage because these can be estimators of two different parameters (see Table 1 above for an example). The conditional odds ratio will generally be farther from 1 than the unconditional odds ratio, and therefore, adjustment for baseline covariates can increase the power of hypothesis testing for superiority despite the increased standard error of treatment effect estimation (Robinson and Jewell 1991).

- Use of nonlinear models such as logistic regression or proportional hazards regression is commonly used in many clinical settings. A semiparametric ordinal regression model (i.e., proportional odds model) can also be used as a flexible method for modeling ordinal or continuous outcomes (Liu et al. 2017).

- 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. 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. Interpretability increases with the quality of model specification.

- Sponsors can perform covariate adjusted estimation and inference for an unconditional treatment effect (e.g., the odds ratio of 4.8 in Table 1) in the primary analysis of data from a randomized trial. The method used should provide valid inference under approximately the same minimal statistical assumptions that would be needed for unadjusted estimation in a randomized trial. If a novel method is proposed and statistical properties are unclear, the specific proposal should be discussed with the review division. Covariate adjusted estimators of unconditional treatment effects that are robust to misspecification of regression models have been proposed for randomized clinical trials with binary outcomes (Ge et al. 2011), ordinal outcomes (Díaz et al. 2016), and time-to-event outcomes (Tangen and Koch 1999); (Lu and Tsiatis 2008).

- The following are steps for one statistically reliable method of covariate adjustment for an unconditional treatment effect with binary outcomes that produces a resulting estimator (Ge et al. 2011); (Freedman 2008) termed “standardized,” “plug-in,” or “g-computation”:

  (1) Fit a logistic model with maximum likelihood that regresses the outcome on treatment assignments and prespecified baseline covariates. The model should include an intercept term.

  (2) For each subject, compute the model-based prediction of the probability of response under treatment in both the treatment group and control group using each subject’s specific baseline covariates.

  (3) Estimate the average response under treatment by averaging (across all subjects in the trial) the probabilities estimated in Step 2.

  (4) For each subject, compute the model-based prediction of the probability of response under control in both the treatment group and control group using each subject’s specific baseline covariates.

  (5) Estimate the average response under control by averaging (across all subjects in the trial) the probabilities estimated in Step 4.

  (6) The estimates of average responses rates in the two treatment groups from Steps 3 and 5 can be used to estimate an unconditional treatment effect, such as the risk difference, relative risk, or odds ratio.

- With nonlinear models using a covariate adjusted estimator for an unconditional treatment effect, sponsors can use the nonparametric bootstrap or standard error formulas justified in the statistical literature for confidence interval construction and hypothesis testing.

IV. REFERENCES

Agresti A, 2002, Categorical Data Analysis, Second Edition. Wiley Online Library.

Bugni F, Canay IA, and AM Shaikh, 2018, Inference Under Covariate-Adaptive Randomization, Journal of the American Statistical Association, 113(524):1784-1796.

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.

Gail MH, Wieand S, and S Piantadosi, 1984, Biased Estimates of Treatment Effect in Randomized Experiments with Nonlinear Regressions and Omitted Covariates, Biometrika, 71(3):431-444.

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

Groenwold RHH, White, IR, Donders, AR, Carpenter JR, Altman DG, and KGM Moons, 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.

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

Liu Q, Shepherd BE, Li C, and FE Harrell, 2017, Modeling Continuous Response Variables Using Ordinal Regression, Statistics in Medicine, 36(27):4316-4335.

Lu X and AA Tsiatis, 2008, Improving the Efficiency of the Log-Rank Test Using Auxiliary Covariates, Biometrika, 95(3):679-694.

Mantel N and W Haenszel, 1959, Statistical Aspects of the Analysis of Data from Retrospective Studies of Disease, Journal of the National Cancer Institute, 22(4):719-748.

Robinson LD and NP Jewell, 1991, Some Surprising Results About Covariate Adjustment in Logistic Regression Models, International Statistical Review, 58(2):227-240.

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.

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.