Covariate adjustment in RCTs

Translating fundamental principles into practice

Dominic Magirr

10th August 2026

Acknowledgement: Mark Baillie, Xinlei Deng, Tim Morris, Alex Pryzbylski, Craig Wang, Cong Zhang

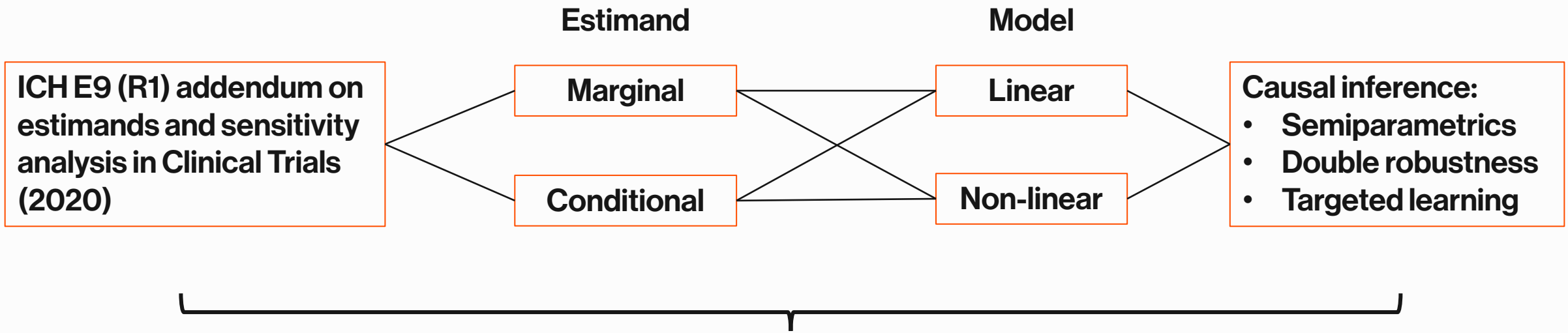
NOVARTIS

Covariate adjustment has been well accepted for a long time

ICH E9 Statistical Principles for Clinical Trials (1998)

"Pre-trial 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 in order to improve precision and to compensate for any lack of balance between treatment groups."

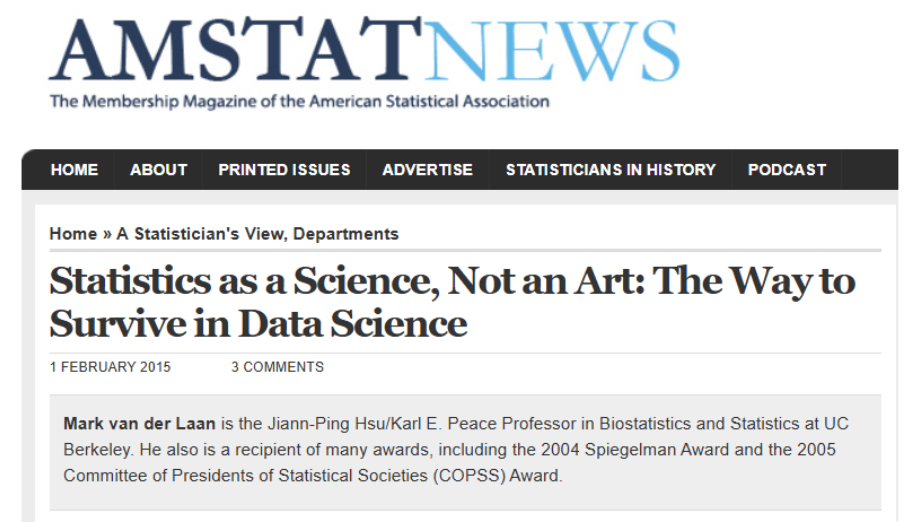
What has changed in recent years?



FDA Guidance for Industry: Adjusting for Covariates in Randomized Clinical Trials (2023)

Movement towards “model-free” estimands and “assumption-lean” analysis methods

The science of statistics



AMSTATNEWS

The Membership Magazine of the American Statistical Association

HOME ABOUT PRINTED ISSUES ADVERTISE STATISTICIANS IN HISTORY PODCAST

Home » A Statistician's View, Departments

Statistics as a Science, Not an Art: The Way to Survive in Data Science

1 FEBRUARY 2015 3 COMMENTS

Mark van der Laan is the Jiann-Ping Hsu/Karl E. Peace Professor in Biostatistics and Statistics at UC Berkeley. He also is a recipient of many awards, including the 2004 Spiegelman Award and the 2005 Committee of Presidents of Statistical Societies (COPSS) Award.

https://magazine.amstat.org/blog/2015/02/01/statscience_feb2015/

The foundation of statistics laid down by its founders [...] could not have been to arbitrarily select a "convenient" statistical model. However, that is precisely what most statisticians blithely do, proudly referring to the quote, "All models are wrong, but some are useful."

[...]

one typically asks a few questions about the data such as: Is the outcome a survival time? Is it case-control data? And then one quickly moves on to returning output from a Cox-Ph model or a logistic regression model with some "reasonable" set of covariates

[...]

Is this mess we have created really necessary? No! As a start, we need to take the field of statistics (i.e., the science of learning from data) seriously. It is complete nonsense to state that all models are wrong, so let's stop using that quote. For example, a statistical model that makes no assumptions is always true.

Model-trusting

$$\text{logit}\, P(Y = 1 \mid A, X) = \alpha_0 + \alpha_1 A + \alpha_2 X$$

- Direct estimation via MLE / posterior probability
- If model is incorrect, it's unclear what α_1 means
- Compatible with Bayesian, likelihood, and frequentist (conditional and unconditional) inference
- Typically used for *conditional estimands*

Model-robust / assumption-lean

$$\overline{Y}_1 - \overline{Y}_0 + \frac{1}{n}\sum_{i=1}^{n}\left(A_i - \frac{1}{2}\right)h(X_i)$$

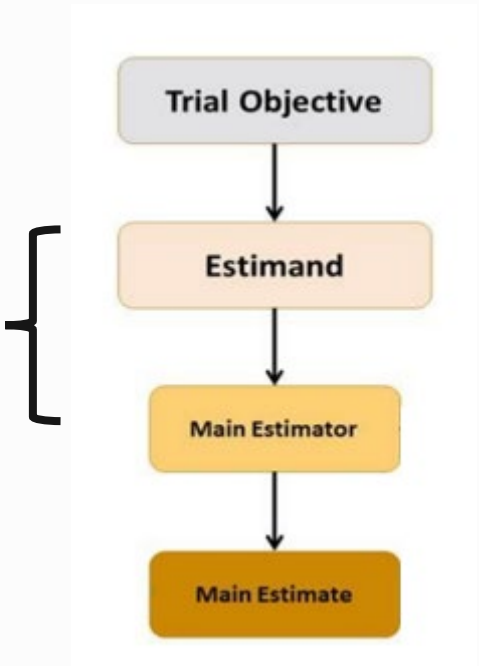
- Combines an unadjusted estimator with an "estimator of zero"
- Clever choice of h(x) to increase efficiency
- An (unconditional) frequentist approach
- Typically used for *marginal estimands*

Buja et al. (2019); Vansteelandt (2021)

ICH E9 (R1) addendum on estimands (2020)

1. *"An estimand is a precise description of the treatment effect reflecting the clinical question posed by a given clinical trial objective. It summarises **at a population level** what the outcomes would be in the same patients under different treatment conditions being compared."*

2. Estimand before estimation



https://thestatsgeek.com/2023/06/23/is-the-ich-e9-estimand-addendum-compatible-with-model-based-estimands/

FDA Guidance on Adjusting for Covariates in RCTs (2023)

*"This guidance is **consistent with the ICH guidance for industry E9(R1)** [...] the treatment effect estimated by covariate adjustment is a population summary measure."*

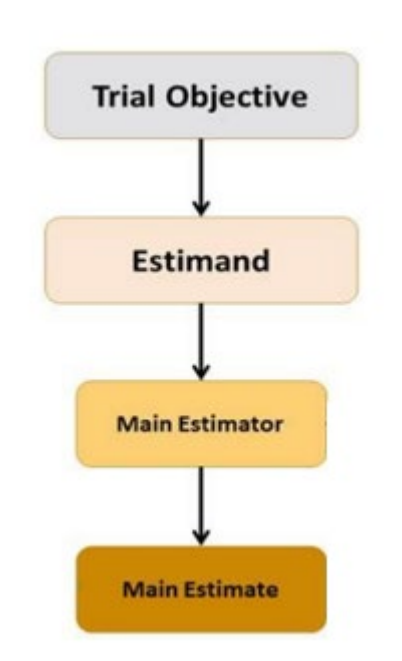
Echoing the philosophy of targeted learning / double robustness:

"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."

Fundamentals: estimands and assumptions

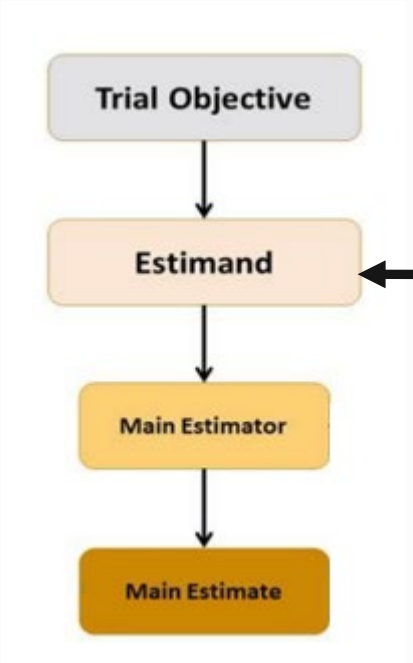
Where to start?

ICH E9 R1 Addendum emphasizes that *estimands* come before *estimators*...



Where to start?

ICH E9 R1 Addendum emphasizes that *estimands* come before *estimators*...

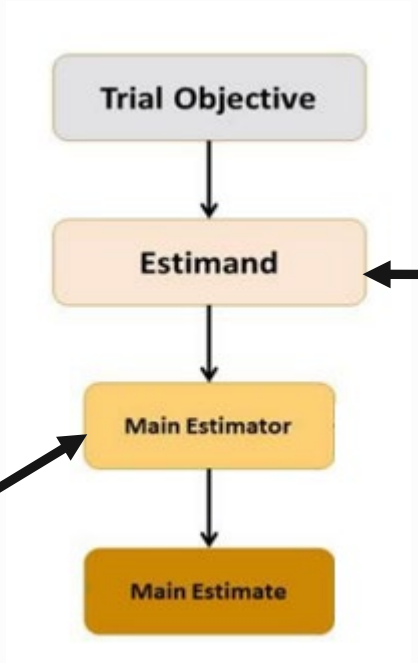


Fundamental modelling assumptions

... but we find it impossible to discuss estimands before at least some fundamental modelling assumptions

Where to start?

ICH E9 R1 Addendum emphasizes that *estimands* come before *estimators*...



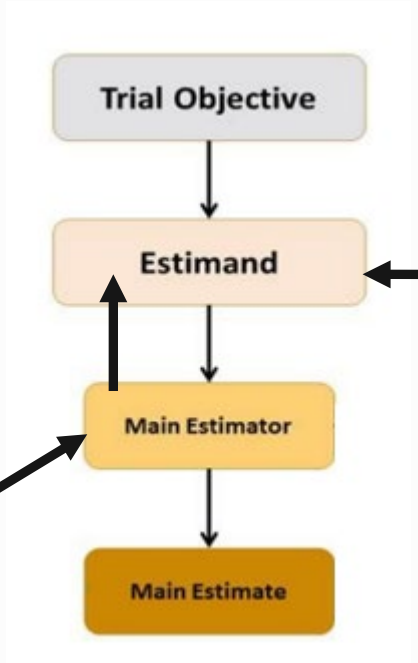
Fundamental modelling assumptions

... but we find it impossible to discuss estimands before at least some fundamental modelling assumptions

Additional modelling assumptions

Where to start?

ICH E9 R1 Addendum emphasizes that *estimands* come before *estimators*...



Fundamental modelling assumptions

... but we find it impossible to discuss estimands before at least some fundamental modelling assumptions

Additional modelling assumptions

Fundamental model assumptions

Option 1: superpopulation inference

$$(Y_i(1), Y_i(0), X_i, A_i) \overset{\text{i.i.d}}{\sim} f, \quad i = 1, \ldots, n$$

$$A \perp\!\!\!\perp Y(1), Y(0), X$$

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

$Y_i(1), Y_i(0)$ are potential outcomes, X_i is a covariate, A_i is assignment

Fundamental model assumptions

Option 1: superpopulation inference

$$(Y_i(1), Y_i(0), X_i, A_i) \overset{\text{i.i.d}}{\sim} f, \quad i = 1, \dots, n$$

$$A \perp\!\!\!\perp Y(1), Y(0), X$$

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

Option 2: randomization inference

$(y_i(1), y_i(0), x_i), i = 1, \dots, n$ are fixed

$$A_i \overset{\text{i.i.d}}{\sim} f$$

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

$Y_i(1), Y_i(0)$ are potential outcomes, X_i is a covariate, A_i is assignment

Fundamental model assumptions

Option 1: superpopulation inference

$$(Y_i(1), Y_i(0), X_i, A_i) \overset{\text{i.i.d}}{\sim} f, \quad i = 1, \ldots, n$$

$$A \perp\!\!\!\perp Y(1), Y(0), X$$

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

Caveats:

- Clinical trials do not use random sampling
- Clinical trials often use stratified block randomization

Option 2: randomization inference

$(y_i(1), y_i(0), x_i), i = 1, \ldots, n$ are fixed

$$A_i \overset{\text{i.i.d}}{\sim} f$$

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

$Y_i(1), Y_i(0)$ are potential outcomes, X_i is a covariate, A_i is assignment

Superpopulation inference

A plethora of estimands

$$(Y_i(1), Y_i(0), X_i, A_i) \overset{\text{i.i.d}}{\sim} f, \quad i = 1, \ldots, n$$

Population average treatment effect, PATE

$$E(Y(1) - Y(0))$$

Superpopulation inference

A plethora of estimands

$$(Y_i(1), Y_i(0), X_i, A_i) \overset{\text{i.i.d}}{\sim} f, \quad i = 1, \ldots, n$$

Population average treatment effect, PATE

$$E(Y(1) - Y(0))$$

More generally,

$$g\{f(Y(1))\} \ \ vs. \ \ g\{f(Y(0))\}$$

Superpopulation inference

A plethora of estimands

$$(Y_i(1), Y_i(0), X_i, A_i) \overset{\text{i.i.d}}{\sim} f, \quad i = 1, \ldots, n$$

Population average treatment effect, PATE

$$E(Y(1) - Y(0))$$

More generally,

$$g\{f(Y(1))\} \ \ vs. \ \ g\{f(Y(0))\}$$

Conditional average treatment effect, CATE(x)

$$E(Y(1) - Y(0) | \, X = x)$$

More generally,

$$g\{f(Y(1) | \, X = x)\} \ \ vs. \ \ g\{f(Y(0) | \, X = x)\}$$

Superpopulation inference

A plethora of estimands

$$(Y_i(1), Y_i(0), X_i, A_i) \overset{\text{i.i.d}}{\sim} f, \quad i = 1, \ldots, n$$

Population average treatment effect, PATE

$$E(Y(1) - Y(0))$$

More generally,

$$g\{f(Y(1))\} \ \ vs. \ \ g\{f(Y(0))\}$$

✓ Easier to estimate with minimal additional assumptions

× May be less relevant to individual decision making

× May not be relevant if target population where the treatment would be applied differs greatly from the patients in the trial

Conditional average treatment effect, CATE(x)

$$E(Y(1) - Y(0) | X = x)$$

More generally,

$$g\{f(Y(1) | X = x)\} \ \ vs. \ \ g\{f(Y(0) | X = x)\}$$

× More difficult to estimate

✓ More relevant for individual decision making

Van Lancker et al. (2024)

FDA Guidance (2023)

*"Sponsors can perform covariate-adjusted estimation and inference for an **unconditional treatment effect** in the primary analysis of data from a randomized trial."*

Vs.

*"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."*

5 Questions we might want to answer in an RCT

Senn (2004)

1. Was there an effect of treatment in this trial?
2. What was the average effect of treatment in this trial?
3. Was the effect identical for all patients in the trial?
4. What was the effect of treatment for different subgroups of patients?
5. What will be the effect of treatment when used more generally (outside the trial)?

5 Questions we might want to answer in an RCT

Senn (2004)

1. Was there an effect of treatment in this trial?
2. What was the average effect of treatment in this trial?
3. Was the effect identical for all patients in the trial?
4. What was the effect of treatment for different subgroups of patients?
5. What will be the effect of treatment when used more generally (outside the trial)?

Scope of *primary* analysis

Unadjusted analysis

What makes unadjusted analysis “acceptable”?

“An unadjusted analysis is acceptable for the primary analysis of an efficacy endpoint”

FDA Guidance Document on Covariate Adjustment (2023)

What makes unadjusted analysis “acceptable”?

“An unadjusted analysis is acceptable for the primary analysis of an efficacy endpoint”

FDA Guidance Document on Covariate Adjustment (2023)

Fundamental assumptions

$$(Y_i(1), Y_i(0), X_i, A_i) \overset{\text{i.i.d}}{\sim} f, \quad i = 1, \ldots, n$$

$$A \perp\!\!\!\perp Y(1), Y(0), X$$

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

$$\Rightarrow \qquad E(\bar{Y}_1 - \bar{Y}_0) = E(Y(1) - Y(0))$$

NOVARTIS

What makes unadjusted analysis “acceptable”?

Model-based ANOVA:

$$\widehat{var}_M := \hat{\sigma}^2 \left(\frac{1}{N_1} + \frac{1}{N_0} \right)$$

Given both the fundamental and working model assumptions, the model-based variance is a consistent estimator

Fundamental assumptions

$$(Y_i(1), Y_i(0), X_i, A_i) \overset{\text{i.i.d}}{\sim} f, \quad i = 1, \ldots, n$$

$$A \perp\!\!\!\perp Y(1), Y(0), X$$

+

Working model assumptions

$$\text{var}(Y_i | A_i) = \sigma^2 \text{ for } i = 1, \ldots, n$$

(Homoskedasticity)

$$n \cdot \widehat{\text{var}}_M \overset{\text{p}}{\rightarrow} n \cdot \text{var}(\bar{Y}_1 - \bar{Y}_0)$$

$\overset{\text{p}}{\rightarrow}$ means convergence in probability

What makes unadjusted analysis "acceptable"?

Model-based ANOVA:

$$\widehat{var}_M := \hat{\sigma}^2 \left(\frac{1}{N_1} + \frac{1}{N_0} \right)$$

Given both the fundamental and working model assumptions, the model-based variance is a consistent estimator

Fundamental assumptions

$$(Y_i(1), Y_i(0), X_i, A_i) \overset{\text{i.i.d}}{\sim} f, \quad i = 1, \ldots, n$$

$$A \perp\!\!\!\perp Y(1), Y(0), X$$

+

Working model assumptions

$$\text{var}(Y_i | A_i) = \sigma^2 \text{ for } i = 1, \ldots, n$$

(Homoskedasticity)

$$n \cdot \widehat{\text{var}}_M \overset{\text{p}}{\rightarrow} n \cdot \text{var}(\bar{Y}_1 - \bar{Y}_0)$$

Sandwich estimator (Eicher-Huber-White):

$$\widehat{var}_S := \frac{\hat{\sigma}_1^2}{N_1} + \frac{\hat{\sigma}_0^2}{N_0}$$

Given only the fundamental assumptions, the sandwich variance is a **robust** and consistent estimator

Fundamental assumptions

$$(Y_i(1), Y_i(0), X_i, A_i) \overset{\text{i.i.d}}{\sim} f, \quad i = 1, \ldots, n$$

$$A \perp\!\!\!\perp Y(1), Y(0), X$$

$$\Rightarrow \quad n \cdot \widehat{\text{var}}_S \overset{\text{p}}{\rightarrow} n \cdot \text{var}(\bar{Y}_1 - \bar{Y}_0)$$

$\overset{\text{p}}{\rightarrow}$ means convergence in probability

Covariate-adjusted analysis

What FDA are looking for from covariate adjusted analysis

Unadjusted analysis achieves the following,

Fundamental assumptions

$$(Y_i(1), Y_i(0), X_i, A_i) \overset{\text{i.i.d}}{\sim} f, \quad i = 1, \ldots, n$$

$$A \perp\!\!\!\perp Y(1), Y(0), X$$

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

$$\Rightarrow$$

$$\hat{\theta} \xrightarrow{p} \theta \qquad (1)$$

$$n \cdot \widehat{\text{var}}\left(\hat{\theta}\right) \xrightarrow{p} n \cdot \text{var}(\hat{\theta}) \qquad (2)$$

(1) Consistent point estimation: as $n \to \infty$, our estimated treatment effect converges to the true treatment effect.

(2) Consistent variance estimation: (unconditional) confidence intervals are 'honest' in the sense that they achieve their advertised coverage level (**at least asymptotically**).

What FDA are looking for from covariate adjusted analysis

Unadjusted analysis achieves the following,

Fundamental assumptions

$$(Y_i(1), Y_i(0), X_i, A_i) \overset{\text{i.i.d}}{\sim} f, \quad i = 1, \ldots, n$$

$$A \perp\!\!\!\perp Y(1), Y(0), X$$

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

$$\Rightarrow$$

$$\hat{\theta} \xrightarrow{p} \theta \qquad (1)$$

$$n \cdot \widehat{\text{var}}\left(\hat{\theta}\right) \xrightarrow{p} n \cdot \text{var}(\hat{\theta}) \qquad (2)$$

(1) Consistent point estimation: as $n \to \infty$, our estimated treatment effect converges to the true treatment effect.

(2) Consistent variance estimation: (unconditional) confidence intervals are 'honest' in the sense that they achieve their advertised coverage level (**at least asymptotically**).

"[Covariate-adjusted analysis] ***should provide valid inference under approximately the same minimal statistical assumptions that would be needed for unadjusted estimation in a randomized trial.****"* – FDA (2023)

What we are looking for from covariate adjusted analysis

Fundamental assumptions

$$(Y_i(1), Y_i(0), X_i, A_i) \overset{\text{i.i.d}}{\sim} f, \quad i = 1, \ldots, n$$

$$A \perp\!\!\!\perp Y(1), Y(0), X$$

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

$$\Rightarrow$$

$$\hat{\theta} \xrightarrow{p} \theta \qquad (1)$$

$$n \cdot \widehat{\text{var}}\left(\hat{\theta}\right) \xrightarrow{p} n \cdot \text{var}(\hat{\theta}) \qquad (2)$$

(1) Consistent point estimation: as $n \to \infty$, our estimated treatment effect converges to the true treatment effect.

(2) Consistent variance estimation: (unconditional) confidence intervals are 'honest' in the sense that they achieve their advertised coverage level (**at least asymptotically**).

What we are looking for from covariate adjusted analysis

Fundamental assumptions

$$(Y_i(1), Y_i(0), X_i, A_i) \overset{\text{i.i.d}}{\sim} f, \quad i = 1, \ldots, n$$

$$A \perp\!\!\!\perp Y(1), Y(0), X$$

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

$\Rightarrow$

$$\hat{\theta} \overset{p}{\to} \theta \qquad (1)$$

$$n \cdot \widehat{\text{var}}\left(\hat{\theta}\right) \overset{p}{\to} n \cdot \text{var}(\hat{\theta}) \qquad (2)$$

$\text{var}(\hat{\theta})$ as small as possible (3)

(1) Consistent point estimation: as $n \to \infty$, our estimated treatment effect converges to the true treatment effect.

(2) Consistent variance estimation: (unconditional) confidence intervals are 'honest' in the sense that they achieve their advertised coverage level (**at least asymptotically**).

(3) We are looking for estimators $\hat{\theta}$ where $\text{var}(\hat{\theta})$ is as small as possible, subject to satisfying (1) and (2).

Linear regression (no interactions)

Consider a linear model,

Working model assumptions

$$E(Y \mid A, X) = \nu_0 + \theta A + \beta X$$

Let $\hat{\theta}$ be the least squares estimator of θ.

Linear regression (no interactions)

Consider a linear model,

Working model assumptions

$$E(Y\,|A, X) = \nu_0 + \theta A + \beta X$$

Let $\hat{\theta}$ be the least squares estimator of θ.

Given **both** the fundamental model assumptions and working model assumptions,

- $E(Y(1) - Y(0)|X = x) = \theta$ for all x θ can be interpreted as a conditional estimand.

Linear regression (no interactions)

Consider a linear model,

Working model assumptions

$$E(Y\ |A, X) = \nu_0 + \theta A + \ \beta X$$

Let $\hat{\theta}$ be the least squares estimator of θ.

Given **both** the fundamental model assumptions and working model assumptions,

- $E(Y(1) - Y(0)|X = x) = \theta$ for all x — θ can be interpreted as a conditional estimand.
- $E\big(Y(1) - Y(0)\big) = \ \theta$ — θ can ALSO be interpreted as a marginal estimand.

Linear regression (no interactions)

Consider a linear model,

Working model assumptions

$$E(Y\,|A,X) = \nu_0 + \theta A + \beta X$$

Let $\hat{\theta}$ be the least squares estimator of θ.

Given **both** the fundamental model assumptions and working model assumptions,

- $E(Y(1) - Y(0)|X = x) = \theta$ for all x — θ can be interpreted as a conditional estimand.
- $E\big(Y(1) - Y(0)\big) = \theta$ — θ can ALSO be interpreted as a marginal estimand.
- $E\big(\hat{\theta}\big) = \theta$ — $\hat{\theta}$ is a consistent estimator of θ.

Linear regression (no interactions)

Consider a linear model,

Working model assumptions

$$E(Y \mid A, X) = \nu_0 + \theta A + \beta X$$

Let $\hat{\theta}$ be the least squares estimator of θ.

Given **only** the fundamental model assumptions,

Linear regression (no interactions)

Consider a linear model,

Working model assumptions

$$E(Y \mid A, X) = \nu_0 + \theta A + \beta X$$

Let $\hat{\theta}$ be the least squares estimator of θ.

Given **only** the fundamental model assumptions,

- $\hat{\theta} \xrightarrow{p} E\big(Y(1) - Y(0)\big)$

$\hat{\theta}$ is a consistent estimator of the marginal, population average treatment effect.

$$\hat{\theta} = \bar{Y}_1 - \bar{Y}_0 - \hat{\beta}(\bar{X}_1 - \bar{X}_0)$$

Consistent for PATE — Consistent for 0

NOVARTIS

Logistic regression

Consider a logistic regression model,

Working model assumptions

$$\text{logit}\{E(Y \mid A, X)\} = \nu + \theta A + \beta X$$

Let $\hat{\theta}$ be the maximum likelihood estimator of θ.

Logistic regression

Consider a logistic regression model,

Working model assumptions

$$\text{logit}\{E(Y\,|A,X)\} = \nu + \theta A + \beta X$$

Let $\hat{\theta}$ be the maximum likelihood estimator of θ.

Given the fundamental model assumptions AND working model assumptions,

- $\text{logit}\{E(Y(1)|X = x)\} - \text{logit}\{E(Y(0)|X = x)\} = \theta$ for all x
- $\hat{\theta} \xrightarrow{p} \theta$

θ can be interpreted as a conditional estimand, and $\hat{\theta}$ is a consistent estimator of θ.

Logistic regression

Consider a logistic regression model,

Working model assumptions

$$\text{logit}\{E(Y\,|A,X)\} = \nu + \theta A + \beta X$$

Let $\hat{\theta}$ be the maximum likelihood estimator of θ.

Given the fundamental model assumptions AND working model assumptions,

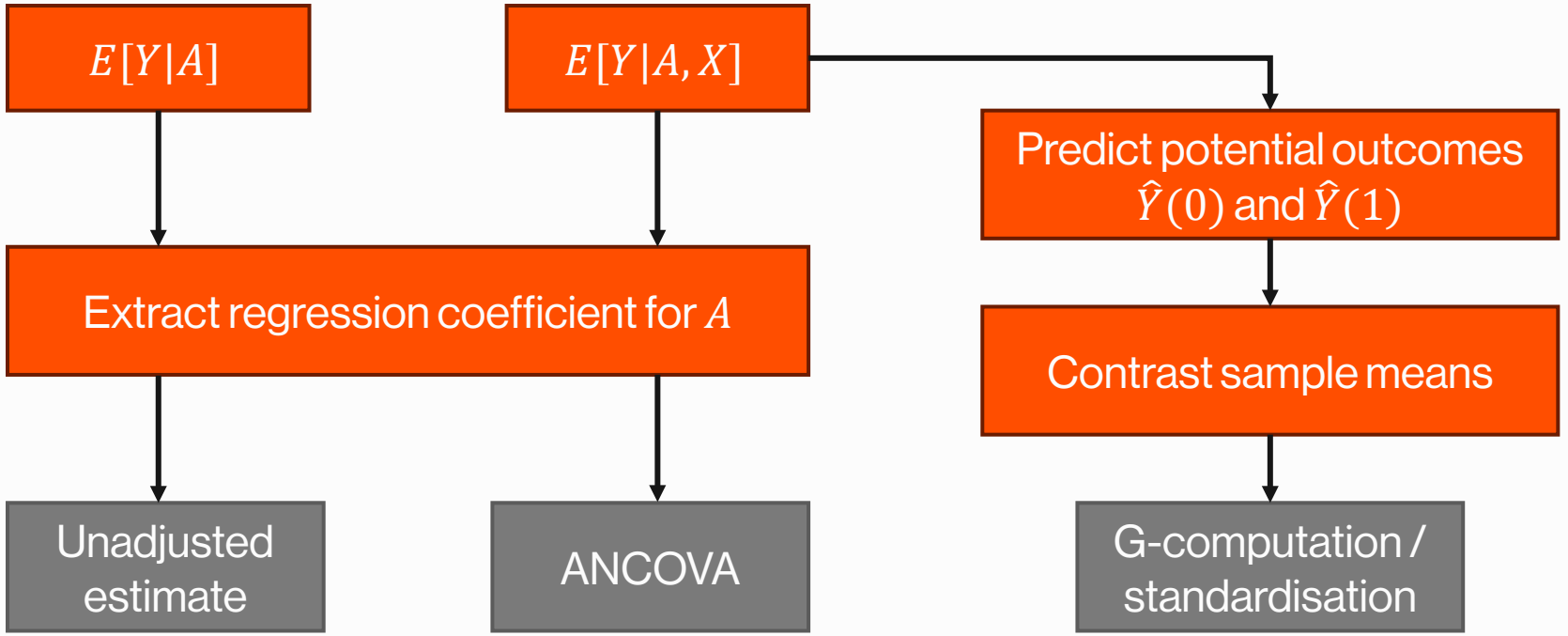
- $\text{logit}\{E(Y(1)|X = x)\} - \text{logit}\{E(Y(0)|X = x)\} = \theta$ for all x
- $\hat{\theta} \xrightarrow{p} \theta$

θ can be interpreted as a conditional estimand, and $\hat{\theta}$ is a consistent estimator of θ.

However, unconditionally

- $\text{logit}\{E(Y(1))\} - \text{logit}\{E(Y(0))\} \neq \theta$ Non-collapsibility

Alternative estimation approach: standardization



Adapted from (Tackney 2023)

Standardization for marginal effect estimation

- Estimate $E[Y(a)]$ using

$$\hat{\pi}_a = \frac{1}{n}\sum_{i=1}^{n} \text{logit}^{-1}\{\hat{v} + \hat{\theta}a + \hat{\beta}X_i\}$$

Standardization for marginal effect estimation

- Estimate $E[Y(a)]$ using

$$\hat{\pi}_a = \frac{1}{n}\sum_{i=1}^{n} \text{logit}^{-1}\{\hat{v} + \hat{\theta}a + \hat{\beta}X_i\}$$

$E[Y(1)] - E[Y(0)]$	$\hat{\pi}_1 - \hat{\pi}_0$
$E[Y(1)] \,/\, E[Y(0)]$	$\hat{\pi}_1 / \hat{\pi}_0$
$\text{logit}\{E[Y(1)]\} - \text{logit}\{E[Y(0)]\}$	$\text{logit}\{\hat{\pi}_1\} - \text{logit}\{\hat{\pi}_0\}$

Why is $\hat{\pi}_a$ consistent for $E[Y(a)]$?

Using only the fundamental model assumptions

Even if our model is wrong, it will converge* to some limit

$$\hat{\pi}_a = \frac{1}{n}\sum_{i=1}^{n} \text{logit}^{-1}\{\hat{\nu} + \hat{\theta}a + \hat{\beta}X_i\}$$

$$\xrightarrow{p} \quad \frac{1}{n}\sum_{i=1}^{n} \text{logit}^{-1}\{\nu^* + \theta^* a + \beta^* X_i\} \quad =: \quad \pi_a^*$$

*convergence as sample size increases

Why is $\hat{\pi}_a$ consistent for $E[Y(a)]$?

Using only the fundamental model assumptions

Even if our model is wrong, it will converge* to some limit

$$\hat{\pi}_a = \frac{1}{n}\sum_{i=1}^{n} \text{logit}^{-1}\{\hat{\nu} + \hat{\theta}a + \hat{\beta}X_i\}$$

$$\xrightarrow{p} \quad \frac{1}{n}\sum_{i=1}^{n} \text{logit}^{-1}\{\nu^* + \theta^* a + \beta^* X_i\} \quad =: \quad \pi_a^*$$

Under randomization, predictions for subjects only on treatment a will converge* to the *same* limit

$$\hat{\pi}_{a,a} = \frac{1}{N_a}\sum_{i=1}^{n} I(A_i = a)\text{logit}^{-1}\{\hat{\nu} + \hat{\theta}a + \hat{\beta}X_i\}$$

$$\xrightarrow{p} \quad \pi_a^*$$

*convergence as sample size increases

Why is $\hat{\pi}_a$ consistent for $E[Y(a)]$?

Using only the fundamental model assumptions

Even if our model is wrong, it will converge* to some limit

$$\hat{\pi}_a = \frac{1}{n}\sum_{i=1}^{n} \text{logit}^{-1}\{\hat{\nu} + \hat{\theta}a + \hat{\beta}X_i\}$$

$$\xrightarrow{p} \quad \frac{1}{n}\sum_{i=1}^{n} \text{logit}^{-1}\{\nu^* + \theta^* a + \beta^* X_i\} \quad =: \quad \pi_a^*$$

Under randomization, predictions for subjects only on treatment a will converge* to the *same* limit

$$\hat{\pi}_{a,a} = \frac{1}{N_a}\sum_{i=1}^{n} I(A_i = a)\text{logit}^{-1}\{\hat{\nu} + \hat{\theta}a + \hat{\beta}X_i\}$$

$$\xrightarrow{p} \quad \pi_a^*$$

We know $\bar{Y}_a \xrightarrow{p} E[Y(a)]$

$$\hat{\pi}_a \; - \; (\,\hat{\pi}_{a,a} \; - \; \bar{Y}_a\,) \; \xrightarrow{p} \; E[Y(a)]$$

*convergence as sample size increases

Why is $\hat{\pi}_a$ consistent for $E[Y(a)]$?

Using only the fundamental model assumptions

Even if our model is wrong, it will converge* to some limit

$$\hat{\pi}_a = \frac{1}{n}\sum_{i=1}^{n} \text{logit}^{-1}\{\hat{\nu} + \hat{\theta}a + \hat{\beta}X_i\}$$

$$\xrightarrow{p} \quad \frac{1}{n}\sum_{i=1}^{n} \text{logit}^{-1}\{\nu^* + \theta^* a + \beta^* X_i\} \quad =: \quad \pi_a^*$$

Under randomization, predictions for subjects only on treatment a will converge* to the *same* limit

$$\hat{\pi}_{a,a} = \frac{1}{N_a}\sum_{i=1}^{n} I(A_i = a)\text{logit}^{-1}\{\hat{\nu} + \hat{\theta}a + \hat{\beta}X_i\}$$

$$\xrightarrow{p} \quad \pi_a^*$$

We know $\bar{Y}_a \xrightarrow{p} E[Y(a)]$

$$\hat{\pi}_a \quad - \quad (\,\hat{\pi}_{a,a} \quad - \quad \bar{Y}_a\,) \quad \xrightarrow{p} \quad E[Y(a)]$$

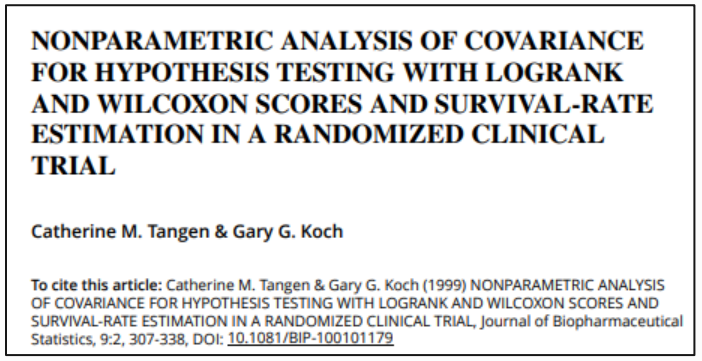
$= 0$ (prediction unbiasedness property of GLM with canonical link)

*convergence as sample size increases

What about time-to-event outcomes?

FDA guidance

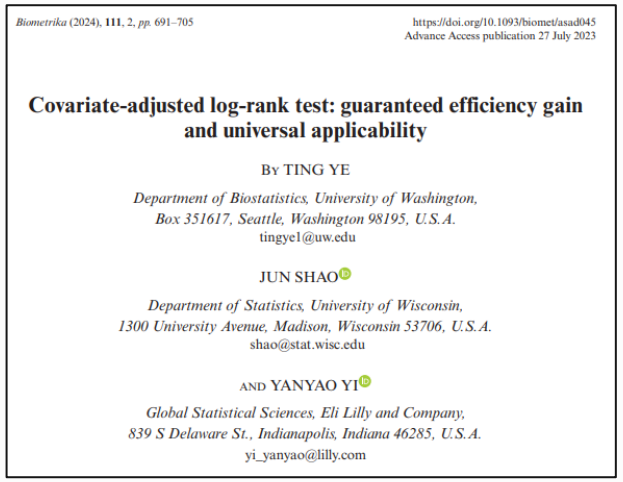
- Covariate-adjusted estimators of unconditional treatment effects that are robust to misspecification of regression models have been proposed for randomized clinical trials with binary outcomes (e.g., Steingrimsson et al. 2017), ordinal outcomes (e.g., Díaz et al. 2016), count outcomes (e.g., Rosenblum and van der Laan 2010), and time-to-event outcomes (e.g., Tangen and Koch 1999; Lu and Tsiatis 2008). If a novel method is proposed and statistical properties are unclear, the specific proposal should be discussed with the review division.



NONPARAMETRIC ANALYSIS OF COVARIANCE FOR HYPOTHESIS TESTING WITH LOGRANK AND WILCOXON SCORES AND SURVIVAL-RATE ESTIMATION IN A RANDOMIZED CLINICAL TRIAL

Catherine M. Tangen & Gary G. Koch

To cite this article: Catherine M. Tangen & Gary G. Koch (1999) NONPARAMETRIC 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, DOI: 10.1081/BIP-100101179



Biometrika (2024), **111**, 2, *pp.* 691–705 https://doi.org/10.1093/biomet/asad045
Advance Access publication 27 July 2023

Covariate-adjusted log-rank test: guaranteed efficiency gain and universal applicability

BY TING YE

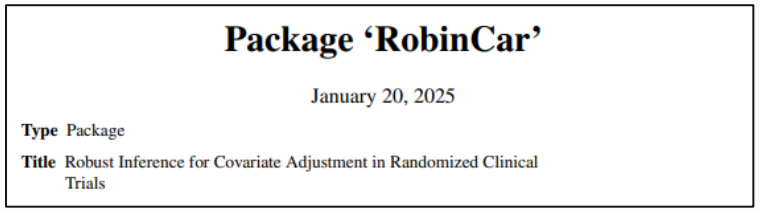
Department of Biostatistics, University of Washington, Box 351617, Seattle, Washington 98195, U.S.A.
tingye1@uw.edu

JUN SHAO

Department of Statistics, University of Wisconsin, 1300 University Avenue, Madison, Wisconsin 53706, U.S.A.
shao@stat.wisc.edu

AND YANYAO YI

Global Statistical Sciences, Eli Lilly and Company, 839 S Delaware St., Indianapolis, Indiana 46285, U.S.A.
yi_yanyao@lilly.com



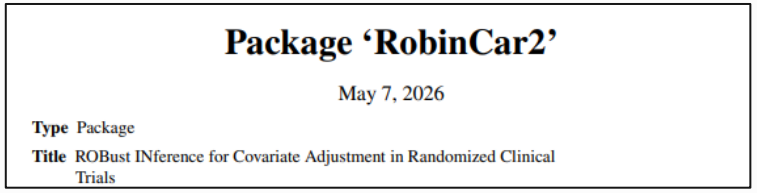
Package 'RobinCar'

January 20, 2025

Type Package

Title Robust Inference for Covariate Adjustment in Randomized Clinical Trials

https://marlenabannick.com/RobinCar/index.html



Package 'RobinCar2'

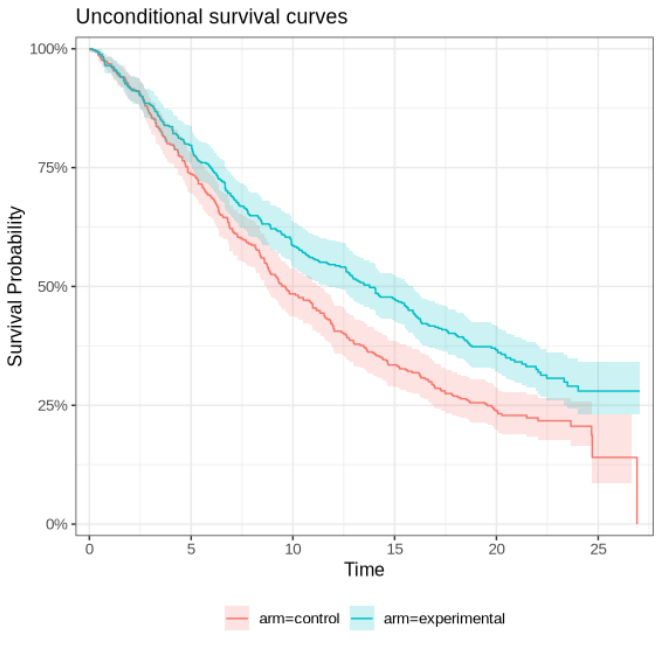
May 7, 2026

Type Package

Title ROBust INference for Covariate Adjustment in Randomized Clinical Trials

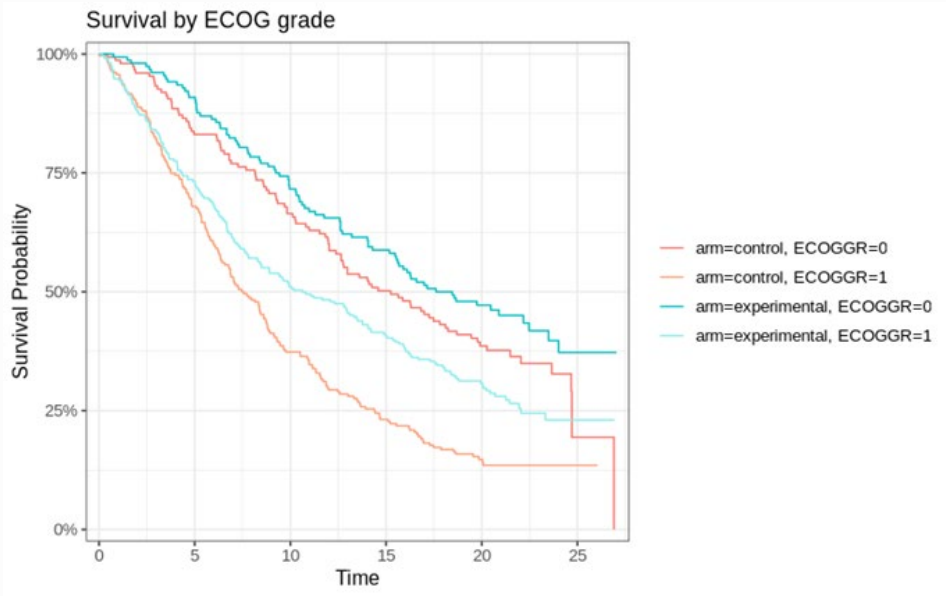
https://openpharma.github.io/RobinCar2/main/

Covariate-adjusted estimator of unconditional hazard ratio



$$S_1(t) = S_0(t)^{\theta_M}$$

$$\hat{\theta}_M = 0.729;\ \widehat{se}\left(\log(\hat{\theta}_M)\right) = 0.0842;\ Z = -3.76$$

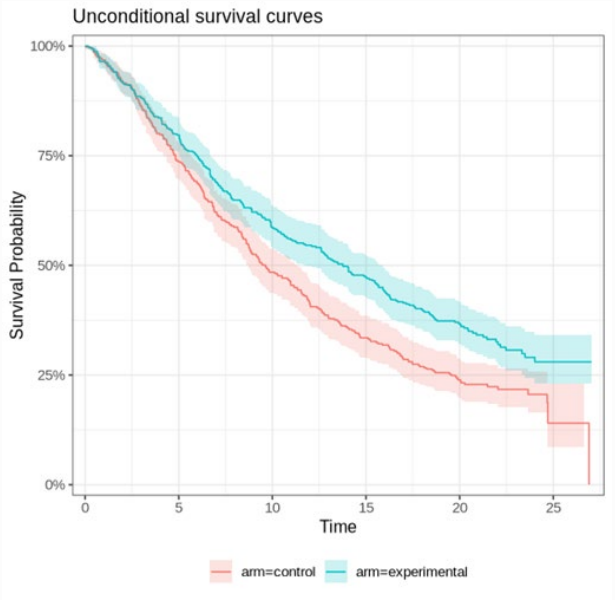


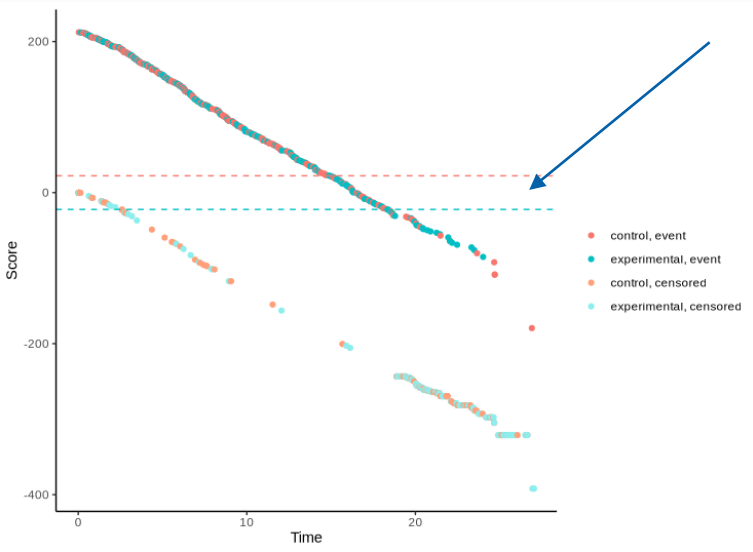
$$S_1(t) = S_0(t)^{\theta_M}$$

$$\hat{\theta}_{M,adj} = 0.724;\ \widehat{se}(\log(\hat{\theta}_{M,adj})) = 0.0816;\ Z_{adj} = -3.97$$

How does this work for hypothesis testing?

$$S_1(t) = S_0(t)^{\theta_M}$$





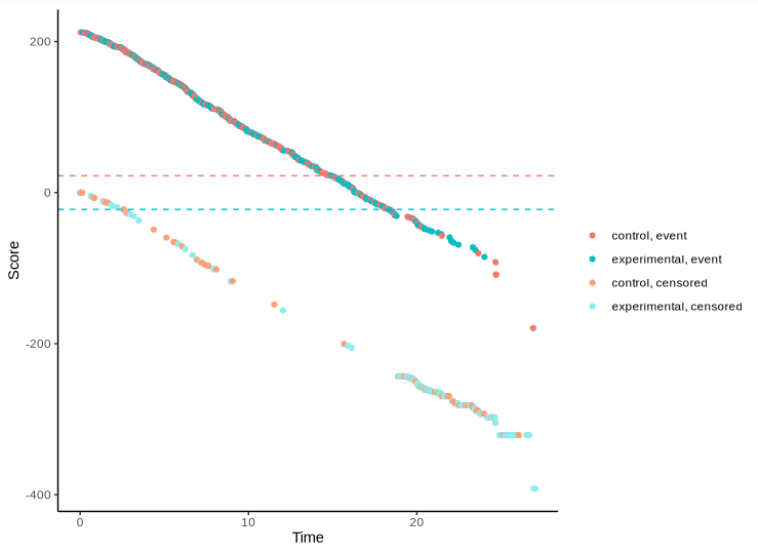
Difference in average score is $Logrank$

$$Logrank = \sum_{t_j} O_{1,j} - E_{1,j} = \cdots = \frac{\sum_{i=1}^{n} A_i Score_i}{\sum_{i=1}^{n} A_i} - \frac{\sum_{i=1}^{n} (1 - A_i) Score_i}{\sum_{i=1}^{n} (1 - A_i)}$$

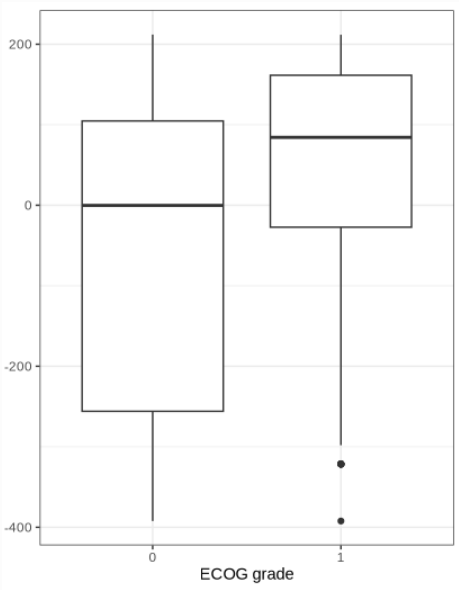
Expected #events on trt 1 under H_0: $\theta_M = 1$

See, e.g., Leton & Zuluaga (2001)

How does this work for hypothesis testing?



Can baseline covariates explain any of the variation in scores?



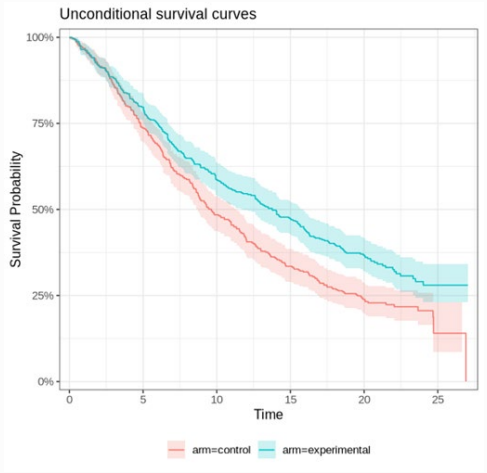
Suggests a linear model: $\mathrm{E}(Score_i|A_i, ECOG_i) = \gamma + \gamma_1 A_i + \gamma_2 A_i(ECOG_i - \overline{ECOG}) + \gamma_3(1 - A_i)(ECOG_i - \overline{ECOG})$

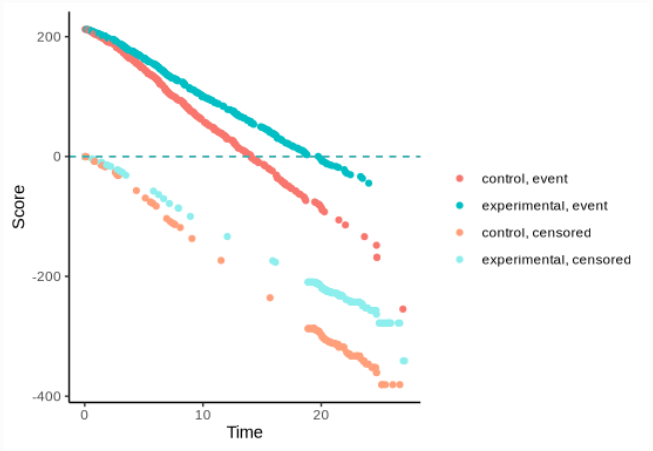
If $|\hat{\gamma_1}|$ large enough, reject $H_0: \theta_M = 1$

$$\hat{\gamma_1} = LR - \{\hat{\gamma}_2(\overline{ECOG}_{trt=1} - \overline{ECOG}) - \hat{\gamma}_3(\overline{ECOG}_{trt=0} - \overline{ECOG})\}$$

How does this work for point estimation?

$$S_1(t) = S_0(t)^{\theta_M}$$





$$\sum_{t_j} O_{1,j} - E_{1,j}(\theta_M^*) = \cdots = \frac{\sum_{i=1}^{n} A_i Score_i(\theta_M^*)}{\sum_{i=1}^{n} A_i} - \frac{\sum_{i=1}^{n}(1-A_i) Score_i(\theta_M^*)}{\sum_{i=1}^{n}(1-A_i)}$$

Unadjusted: find θ_M^* such that

$$LR(\theta_M^*) = 0$$

Adjusted: find θ_M^* such that:

$$LR(\theta_M^*) - \{\hat{\gamma}_2(\overline{ECOG}_{trt=1} - \overline{ECOG}) - \hat{\gamma}_3(\overline{ECOG}_{trt=0} - \overline{ECOG})\} = 0$$

Implementation

- Ye et al. (2024): refinement of Tang & Koch (1999) + asymptotic theory
- Bannick et al. (2024): implementation in {RobinCar}
- ASA-BIOP Covariate Adjustment WG: {RobinCar2}
  - Simple syntax

```
RobinCar2::robin_surv(
  Surv(time, event) ~ ECOGGR + blSLD,
  treatment = arm ~ 1,
  data = dat
)
```

Key opportunity: adjust for continuous covariates such as baseline tumour size (or also supercovariates) without changing the target estimand, or sacrificing robustness.

Note on variance estimation for covariate-adjusted estimators

Same assumptions as an unadjusted analysis

$\Rightarrow$

$$\hat{\theta} \xrightarrow{p} \theta \quad (1)$$

$$n \cdot \widehat{\mathrm{var}}\left(\hat{\theta}\right) \xrightarrow{p} n \cdot \mathrm{var}(\hat{\theta}) \quad (2)$$

$\mathrm{var}(\hat{\theta})$ as small as possible (3)

- We have focused on (1) .
- Variance estimators exist that satisfy (2) for continuous, binary, time-to-event outcomes...
  - But often rely on asymptotic arguments.
  - We must consider finite-sample properties:
    - Linear models – see Senn, Koenig & Posch (2026) https://doi.org/10.1177/09622802261454833
    - Time-to-event – see Krotka & Magirr (2026) https://doi.org/10.48550/arXiv.2608.10923

Summary

- Covariate adjustment has been used in primary analysis of clinical trials for a long time.
  - Recent extra scrutiny on robustness and "estimand before estimation" .

- Defining an estimand without any assumptions is not possible. We need some kind of mathematical framework to make progress. Superpopulation framework is a pragmatic flexible choice. Important caveats:
  - RCTs do not use random samples – we need to think carefully about what the superpopulation means.
  - RCTs often use stratified block randomization.

- Conditional estimands are typically more relevant than marginal estimands for patient-level decision making.
  - But they are simply more difficult to estimate without strong modelling assumptions.
  - That's why we've focused on marginal estimands as the building block for RCT *primary* analysis.

- Unadjusted analysis in an RCT is considered "valid" according to FDA, based on its unconditional properties.
  - Accordingly, we've focused on unconditional properties of covariate-adjusted estimators.

- Our aim: make $var(\hat{\theta})$ as small as possible, by adjusting for the most prognostic baseline covariates, whilst maintaining "valid" inference in the sense of the FDA guidance. We do not change the estimand. We do not sacrifice robustness.

- We can do this with modern methods (standardization / g-computation and the covariate-adjusted hazard ratio of *Ye et al., 2024*)

References

van der Laan, Mark. "Statistics as a science, not an art: the way to survive in data science." *Amstat News* 1 (2015): 292.

Buja, Andreas, et al. "Models as approximations I." *Statistical Science* 34.4 (2019): 523-544.

Vansteelandt, Stijn. "Statistical modelling in the age of data science." *Observational Studies* 7.1 (2021): 217-228

Van Lancker, Kelly, Frank Bretz, and Oliver Dukes. "Covariate adjustment in randomized controlled trials: General concepts and practical considerations." *Clinical Trials* 21.4 (2024): 399-411.

Senn, Stephen. "Controversies concerning randomization and additivity in clinical trials." *Statistics in medicine* 23.24 (2004): 3729-3753.

ICH, Statistical Principles for Clinical Trials E9 (International Council for Harmonisation, 1998).

FDA, Adjusting for Covariates in Randomized Clinical Trials for Drugs and Biological Products. Guideline for Industry (FDA, 2023).

ICH, ICH E9 (R1) Addendum on Estimands and Sensitivity Analysis in Clinical Trials to the Guideline on Statistical Principles for Clinical Trials (ICH, 2020).

Tackney, Mia S., et al. "A comparison of covariate adjustment approaches under model misspecification in individually randomized trials." *Trials* 24.1 (2023): 14.

Ye, T., Shao, J., & Yi, Y. (2024). Covariate-adjusted log-rank test: guaranteed efficiency gain and universal applicability. *Biometrika*, *111*(2), 691-705.

Tangen, C. M., & Koch, G. G. (1999). Nonparametric 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.

Gandara, D. R., Paul, S. M., Kowanetz, M., Schleifman, E., Zou, W., Li, Y., ... & Shames, D. S. (2018). Blood-based tumor mutational burden as a predictor of clinical benefit in non-small-cell lung cancer patients treated with atezolizumab. *Nature medicine*, *24*(9), 1441-1448.

Bannick M, Ye T, Yi Y, Bian F (2024). *RobinCar: ROBust INference for Covariate Adjustment in Randomized clinical trials*. R package version 0.3.0.

Senn, Stephen, Franz König, and Martin Posch. "Stratification in small randomised clinical trials and analysis of covariance: Some simple theory and recommendations." Statistical Methods in Medical Research (2026): 09622802261454833.