

Statistical Analysis Plan Appendix

Pre-specified Prognostic Covariate Adjustment

2026-10-27

1 Purpose

This appendix fixes, before unblinding, the model and the single covariate used in the primary analysis of the primary continuous endpoint. It is written to be attached to the statistical analysis plan and to be read by a reviewer who wants to check that the adjustment was specified in advance rather than chosen after the data were seen.

It does not report the observed treatment effect. The design quantities in Section 6 are inputs fixed from the source dataset named in Section 9, not results.

2 Estimand

The estimand is the average treatment effect on the primary continuous endpoint in the randomized population,

$$\tau = E[Y(1) - Y(0)],$$

with treatment assigned at randomization. The analysis is intention to treat. The randomization ratio in the source dataset is 0.5000 treated to 0.5000 control.

3 Prognostic score

The covariate is a scalar prognostic score $f(X)$, fixed before unblinding. It predicts the control arm outcome from baseline covariates only, and it is constructed in one of two ways, chosen before the trial data are used:

1. **External fit.** The score is fitted on a cohort whose outcomes are independent of the trial outcomes, then applied unchanged to the trial.
2. **Cross-fitting.** The score is fitted inside the trial by K -fold cross-fitting, and each patient is scored by a model that did not see that patient's outcome.

In both cases the trial outcome of a patient never contributes to the prediction used for that patient. The score enters the primary model as a single covariate. No further covariate selection is performed in the analysis.

4 Primary analysis model

The primary analysis is a linear model with treatment and the prognostic score,

$$Y_i = \alpha + \tau T_i + \beta f(X_i) + \varepsilon_i \quad (1)$$

where τ is the estimand. The treatment coefficient is estimated by least squares, with a heteroskedasticity-consistent (HC0) standard error, and inference uses a two-sided normal interval at level $1 - \alpha$.

The same estimator has a prediction-powered form. With one tuning value shared between arms, the external cohort cancels algebraically and

$$\hat{\tau} = \bar{Y}_1 - \bar{Y}_0 - \hat{\lambda} (\bar{f}_1 - \bar{f}_0) \quad (2)$$

where $\bar{Y}_w$ and $\bar{f}_w$ are arm means. Equations (2) and (3) are the same estimator as (1) when the arms are balanced: the tuning value is the pooled within-arm slope of the outcome on the score, which is the regression coefficient β after partialling out treatment.

5 Tuning

The tuning value is estimated from the trial data alone by

$$\hat{\lambda} = \frac{\widehat{\text{Cov}}_1(Y, f)/n_1 + \widehat{\text{Cov}}_0(Y, f)/n_0}{\widehat{\text{Var}}_1(f)/n_1 + \widehat{\text{Var}}_0(f)/n_0} \quad (3)$$

with biased (divide by n) covariance and variance so that (2) reproduces the least squares coefficient exactly at equal allocation. No unlabeled cohort quantity enters (2). The interval uses

$$\widehat{\text{Var}}(\hat{\tau}) = \frac{\widehat{\text{Var}}_1(Y - \hat{\lambda}f)}{n_1} + \frac{\widehat{\text{Var}}_0(Y - \hat{\lambda}f)}{n_0} \quad (4)$$

Because $\hat{\lambda}$ is a scalar estimated from the same data, its contribution to the variance is of smaller order than the terms in (4), which treat it as fixed.

If a per-arm form with arm-specific tuning values is used as a sensitivity analysis, the external cohort no longer cancels. That form is unbiased only if the external cohort is exchangeable with the trial population in the covariates that enter the score. The per-arm form is not the primary analysis.

6 Design and sample size

Under the prognostic covariate adjustment model with balanced allocation, equal arm variances σ^2 and prognostic correlation ρ , the per-patient variance constant and the required total sample size are

$$v = 4\sigma^2 (1 - \rho^2), \quad n = \frac{(z_{1-\alpha/2} + z_{\text{power}})^2 v}{\Delta^2} \quad (5)$$

The design inputs below are used for planning. The effect and the outcome standard deviation are taken from the source dataset; the prognostic correlation is 0.3099 in squared form, from the source named in the sources column. These are projections under the stated model. They are not observed savings from a completed trial.

Design input	Value
Two-sided alpha	0.050
Target power	80%
Assumed effect Δ	0.2361
Outcome standard deviation σ	1.2295
Squared prognostic correlation ρ^2	0.3099
Per-patient variance, unadjusted	6.0466
Per-patient variance, adjusted	4.1725
Required total n , unadjusted	852
Required total n , adjusted	588
Patients saved at the target power	264
Planned trial size	720
Interval width ratio	0.8307

At the planned size of 720 patients, the adjusted analysis carries the information of approximately 1043 patients under an unadjusted analysis at the same effect and variance.

7 Missing data and sensitivity analyses

The primary analysis is complete case. No imputation is used, and the number of excluded rows is reported. Pre-specified sensitivity analyses are: the unadjusted difference in means; adjustment for the score with the per-arm prediction-powered form where an external cohort is available; and, where the score was cross-fitted, a repeat with a different fold assignment.

8 Reporting

The report gives the point estimate of τ , its standard error, the two-sided interval at level $1 - \alpha$, and the corresponding p value. Subgroup analyses are not part of this appendix and are not claimed to be adjusted by this model.

9 Software and provenance

The analysis was produced by `ppict` version 0.1.0. The dataset is identified by its SHA-256 digest, so the exact input can be checked even if the file path is temporary. The command that produced the numbers in this appendix was:

```
command: python -m ppict.cli analyze
        --csv calculator/data/simulated_trial.csv
        --cost 4200
        --power 0.8
        --patients-per-month 25
dataset: calculator/data/simulated_trial.csv
dataset sha256: 90d62a6ee95c48fc50c5c89288ab4ac18a905ecfb1bf4d33045426096cf56844
trial rows: 720 (360 treated, 360 control)
external rows: 600
ppict version: 0.1.0
alpha: 0.05
```