

Application of PTE methods to a Hypothetical Vaccine Trial

Florian Stijven

2023-09-01

Introduction

In this document, the analysis of a set of hypothetical vaccine trials is described. This document supplements the review of the proportion of treatment effect explained (PTE) as a measure of surrogacy. All data in this document are hypothetical, but meant to be realistic. The goal of this document is **not** to show how different versions of the PTE are estimated in practice. Instead, the goal is to illustrate three conceptual points related to the PTE:

1. The PTE can be interpreted in a variety of ways. However, for the PTE to have a relevant interpretation, unverifiable assumptions are **always** needed.
2. The PTE is often defined independent of baseline covariates. Without taking such baseline covariates into account, the PTE can often not be interpreted in the desired manner.
3. The validity of the non-causal interpretation of the PTE is inherently linked to the potential application trials. The validity of the causal interpretation does not involve assumptions regarding potential application trials.

In this hypothetical scenario, we have data from a single evaluation trial in which the efficacy of a vaccine is evaluated. In this trial, the occurrence of infection in the 120 days after vaccination is the primary endpoint, i.e., the true endpoint, T . At the same time, the antibody levels 2 weeks after vaccination have been measured. This serves as a potential surrogate endpoint, S .

Independent of the evaluation trial, we consider two potential application trials. These are trials in which the results of the previous surrogacy analysis are used to replace the true endpoint with the surrogate endpoint. The population in these two application trials differs from the evaluation trial's population. In addition, the vaccine in the second application trial has a different mechanism of action. These differences are meant to emulate a real-life setting where, indeed, there may be important differences between trials for the same disease. This also puts the spotlight on the reasoning required for justifying the use of a potential surrogate endpoint in a new trial, and the potential pitfalls.

The remainder of the document is structured as follows. First, the data generating model underlying the evaluation trial is explained. For this trial, two versions of the PTE are then computed and interpreted in a causal framework. Next, the data generating model underlying the two potential application trials is explained. The results of the surrogacy evaluation in the evaluation trial are then used to predict the treatment effects in these two application trials *using only surrogacy information*. The accuracy of these predictions is then related to (violations of) assumptions and the PTE's that were estimated in the evaluation trial.

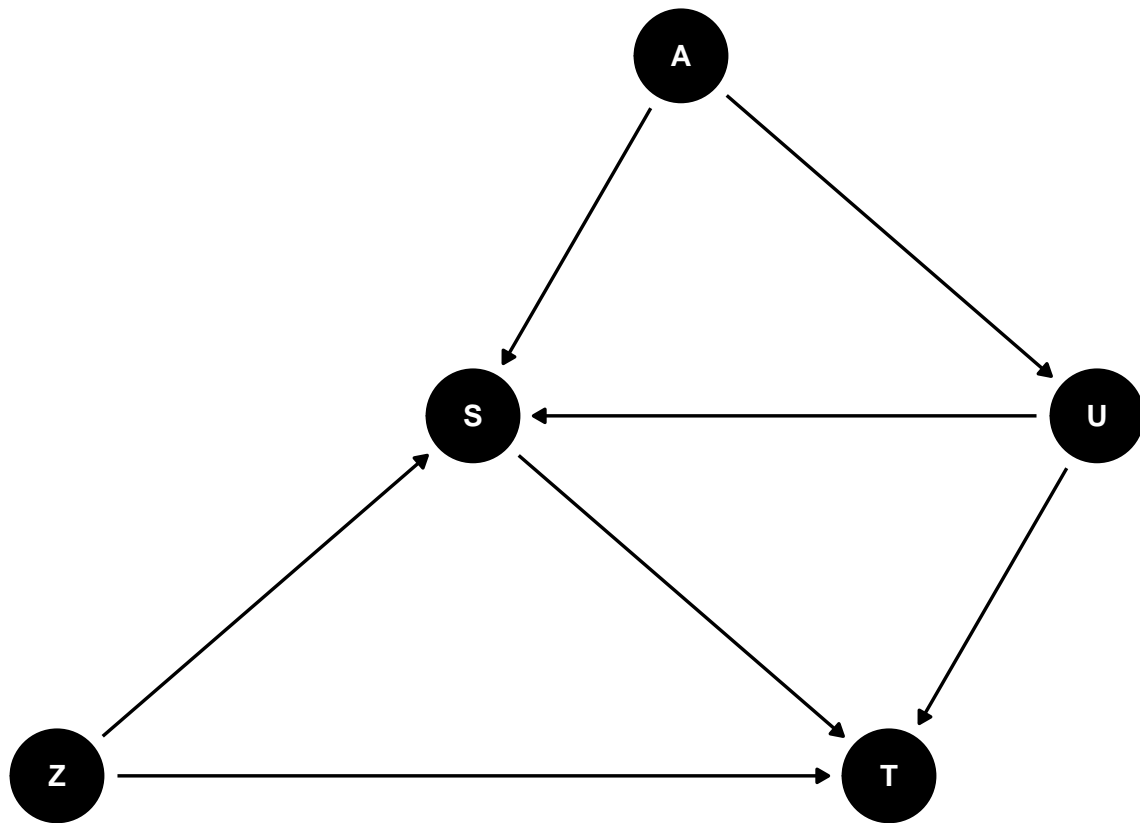
```
knitr::opts_chunk$set(message = FALSE, warning = FALSE)
library(tidyverse)
library(mediation)
# Set a seed for reproducibility.
set.seed(1)
```

Evaluation Trial

Data Generating Model

The data generating model is based on the following selection diagram.

```
coords = list(  
  x = c(Z = 0, S = 0.5, `T` = 1, U = 1.25, A = 0.75),  
  y = c(Z = 0, S = 0.75, `T` = 0, U = 0.75, A = 1.5)  
)  
vaccine_trial_dag = ggdag::dagify(  
  S ~ Z,  
  S ~ U,  
  S ~ A,  
  `T` ~ S,  
  `T` ~ U,  
  `T` ~ Z,  
  U ~ A,  
  coords = coords  
)  
ggdag::ggdag(vaccine_trial_dag) +  
  theme_void()
```



The nodes in the above selection diagram correspond to the following variables:

- **Z.** This is the treatment indicator where $Z = 0$ corresponds to the control vaccine and $Z = 1$ corresponds to the experimental vaccine. Treatment is randomized in this scenario. Therefore, there are no arrows into Z .
- **S.** This is the surrogate endpoint, i.e., the neutralizing antibody level. This variable *causally* affects the true endpoint, i.e., it is a mediator for the treatment effect.
- **T.** This is the true endpoint where $T = 0$ corresponds to infection in 120 day-period following vaccination and $T = 1$ to no infection in this period.
- **U.** This represents the confounding variables for the causal effect of neutralizing antibody levels on the infection outcome. In this data generating model, there are only 2 confounding variables. These variables are measured at baseline.
 - Age, U_1 . Older patients tend to have a weaker immune response to vaccination, and are more susceptible for infection.
 - General health at baseline, U_2 . Patients in better general health tend to have a stronger immune response to vaccination, and are less susceptible for infection.
- **A.** This is the trial indicator where $A = 1$ corresponds to the evaluation trial and $A = 0$ corresponds to the application trial.

We now discuss the distribution of these variables in the evaluation trial in turn. Concomitant with these distribution, we provide the code that is used to generate the data in the evaluation trial.

Treatment

The treatment is randomized in a 1:1 fashion such that there are 10.000 patients in each treatment arm.

```
# Sample size per treatment arm is fixed.
n = 1e4
# Construct integer vector where the first n elements are zero, and the next n
# elements 1.
Z = c(rep(0L, n), rep(1L, n))
```

Age and General Health

As mentioned previously, age and general health are baseline covariates that confound the causal effect of S on T . These variables are distributed in the evaluation trial as follows.

- Age, $U_1|A = 1 \sim N(35, 7^2)$. This population is thus quite young.
- General health, $U_2|A = 1 \sim N(5, 5^2)$. This population has quite good general health (in comparison with the application trials, see further).

Because of randomization, these two variables are independent of the treatment variable.

```
# Generate observations for the baseline covariates.
age = rnorm(n = 2 * n, mean = 35, sd = 7)
health = rnorm(n = 2 * n, mean = 5, sd = 5)

# Combine the treatment vector and baseline covariates into a tibble.
eval_trial_tbl = tibble(
  Z = Z,
  age = age,
  health = health
)
```

Neutralizing Antibody Levels, S

As mentioned before, the vaccine in the evaluation trial mainly exerts its effect through inducing an antibody response. In other words, the vaccine reduces the risk of infection by increasing the patient's neutralizing antibody levels. In the evaluation trial, $S|Z, U, A = 1$ is a normal distribution with unit variance and mean $E(S|Z, U, A = 1) = 1.5 \cdot Z - 0.05 \cdot (U_1 - 45) + 0.15 \cdot U_2$.

```
# The surrogate endpoint is simulated taking into account the baseline
# covariates and treatment assignment.
eval_trial_tbl = eval_trial_tbl %>%
  mutate(S = 1.5 * Z - 0.05 * (age - 45) + 0.15 * health + rnorm(n = 2 * n))
```

Infection Status, T

The true endpoint is a binary endpoint that depends on the treatment, the antibody levels, and the baseline covariates. Its distribution is completely determined by the mean, which depends on the mentioned variables in the following way,

$$E(T|Z, S, U) = \text{expit}(1 + 0.1 \cdot Z + 0.5 \cdot S - 0.03 \cdot (U_1 - 45) + 0.1 \cdot U_2).$$

```
# The true endpoint is simulated in two steps. First, the mean as function of
# covariates is computed. Given this conditional mean, the true endpoint is sampled
# from a bernouilli distribution.
eval_trial_tbl = eval_trial_tbl %>%
  mutate(eta = 0.8 + 0.1 * Z + 0.5 * S - 0.03 * (age - 45) + 0.1 * health,
         infection_free = rbinom(n = 2 * n, size = 1, prob = 1 / (1 + exp(-1 * eta))))

# Add more informative variable names and labels.
eval_trial_tbl = eval_trial_tbl %>%
  mutate(Treatment = factor(Z, levels = 0:1, labels = c("Control", "Experimental")))
```

Surrogacy Analysis

In this subsection, we first do several simple analyses of the evaluation trial that will help us to interpret the results regarding the PTE. The key concepts from the paper are also repeated for completeness. Next, two versions of the PTE are estimated for these data: (i) the “usual” PTE that ignores baseline covariates, PTE_{WT} , and (ii) the PTE that does take into account baseline covariates, PTE_{WT}^X . These estimates are then interpreted in the causal frameworks discussed in the paper.

Preliminaries

We start by computing the marginal treatment effect.

```
eval_trial_prop0 = mean(eval_trial_tbl$infection_free[eval_trial_tbl$Z == 0])
eval_trial_prop1 = mean(eval_trial_tbl$infection_free[eval_trial_tbl$Z == 1])
```

The proportions infection-free in the control and experimental groups are, respectively, 0.862 and 0.932. The risk difference is thus 0.07. In this context, the treatment effect is often summarized in the vaccine efficacy (VE), $VE = 1 - RR$ where RR is the relative risk of infection in the experimental vs control group. The VE in this trial is 0.509.

In what follows, PTE_{WT} is defined as $\frac{\Delta - \Delta_S}{\Delta}$ where

$$\Delta = E(T_1 - T_0 | A = 1)$$

and

$$\Delta_S = \int E(T_1 | S_1 = s, A = 1) dF_{S_0|A}(s|a = 1) - \int E(T_0 | S_0 = s, A = 1) dF_{S_0|A}(s|a = 1).$$

In light of the general definition used throughout the paper, we thus have that $h(u) = u$ and $g(\cdot)$ is the expectation. These developments illustrate that PTE_{WT} will be close to 1 if the regression functions of T on S are similar in both treatment groups. Indeed, we can rewrite Δ_S as the weighted difference of of this two regression functions,

$$\Delta_S = \int E(T_1 | S_1 = s, A = 1) - E(T_0 | S_0 = s, A = 1) dF_{S_0|A}(s|a = 1),$$

where the weights depend on the distribution of the surrogate in the control group. These two regression functions are plotted next.

```
eval_trial_tbl %>%
  ggplot() +
  geom_smooth(aes(x = S, y = infection_free, color = Treatment)) +
  geom_density(aes(x = S, color = Treatment)) +
  xlim(c(-2.5, 7.5)) +
  theme_bw()
```

These regression functions clearly differ between the treatment groups. Indeed, the regression function in the control group is larger than in the experimental group for all values of S . Hence, PTE_{WT} will be larger than 1.

We can also include baseline covariates in the definition of the proportion explained. This leads to $PTE_{WT}^X = \frac{\Delta - E\{\Delta_S(U)\}}{\Delta}$ where

$$\Delta_S(u) = \int E(T_1 | S_1 = s, U = u, A = 1) - E(T_0 | S_0 = s, U = u, A = 1) dF_{S_0|A}(s|a = 1).$$

When U contains all confounders, then $E\{\Delta_S(U)\}$ corresponds to the interventional direct effect and $\Delta - E\{\Delta_S(U)\}$ to the interventional indirect effect. Under the additional cross-world independence assumption, these effects correspond to their natural counterparts. The ratio of the interventional/natural indirect effect and the total effect is classically termed the proportion mediated. As for the PTE, this is actually is misnomer because it is not a proportion.

Estimation of PTE

As explained in Appendix E of the paper, PTE_{WT} can be estimated with existing software for causal mediation analysis. In this analysis, we use the `mediate()` function from the `mediation` package for this purpose. For a conventional mediation analysis, this function relies on two regression models.

1. A regression model for the mediator given the treatment and the confounders measured at baseline.
2. A regression model for the true endpoint given the treatment, the mediator and the confounders measured at baseline.

For estimating the PTE_{WT} , we replace the mediator with the surrogate, and we do not include baseline covariates. For estimating the PTE_{WT}^X , we do include baseline covariates.

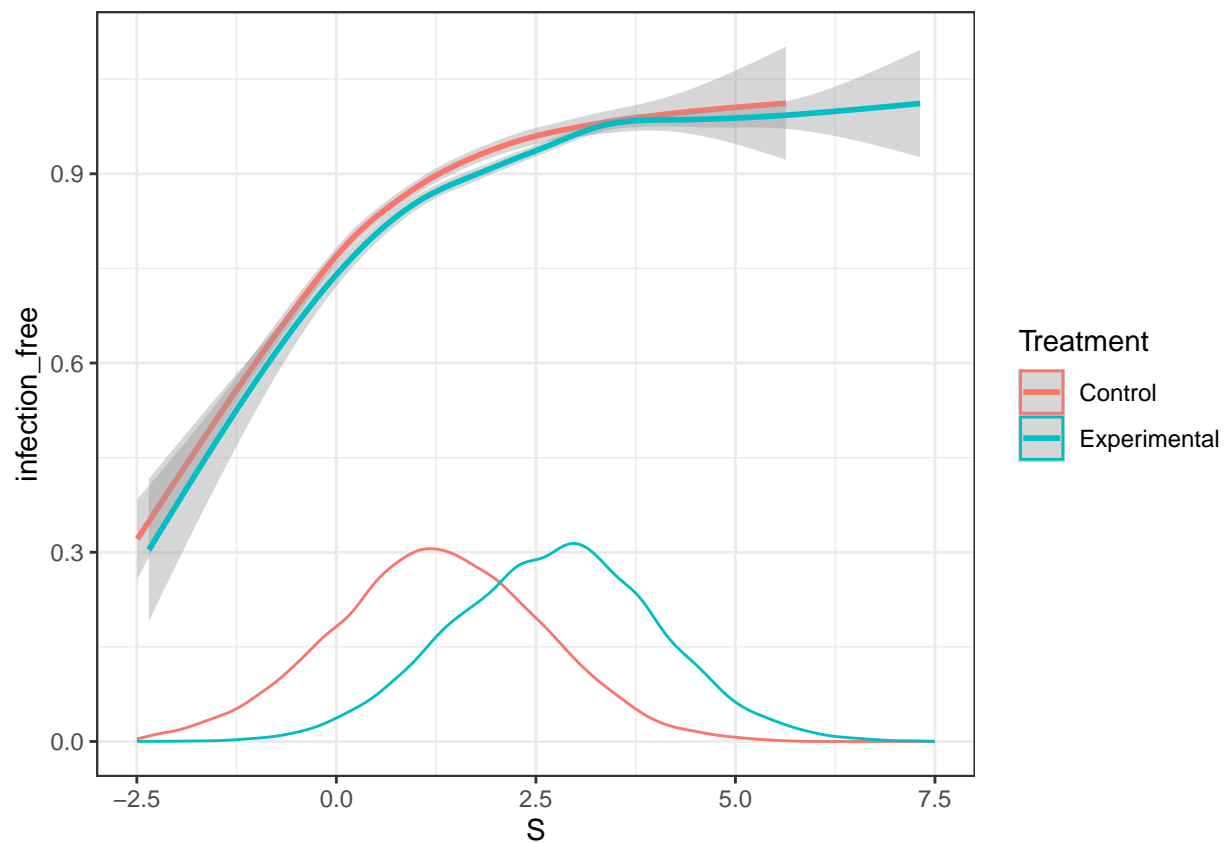


Figure 1: Regression function of the regression of the true endpoint on the surrogate endpoint in each treatment group separately. These regression functions are estimated by local regression. The corresponding density estimates of the distribution of the surrogate endpoint are superimposed on this plot.

The corresponding models are fitted in turn next. For the surrogate, we choose a linear regression model. For the true endpoint, we choose a logistic regression model. When we include the baseline covariates in these models, they are correctly specified. However, the logistic regression model is misspecified when we do not include the baseline covariates.

```
# Fit the models that do not include the baseline covariates.
surrogate_marg_model = lm(S~Z, data = eval_trial_tbl)
true_endpoint_marg_model = glm(infection_free ~ Z + S, data = eval_trial_tbl,
                               family = binomial())

# Fit the models that include the baseline covariates.
surrogate_model = lm(S~Z + age + health, data = eval_trial_tbl)
true_endpoint_model = glm(infection_free ~ Z + S + age + health,
                          data = eval_trial_tbl,
                          family = binomial())
```

We start by estimating PTE_{WT} . For completeness, we call the summary method on the object returned by `mediate()`. This method returns numerous estimates, which are discussed in turn. We only focus on the estimates with `(treated)` in their name. These estimates correspond to components of PTE_{WT} as defined previously. If we switch both treatment arms, then the estimates with `(control)` in their name are obtained.

- Average causal mediation effect, ACME `(treated)`. This is often called the indirect effect. Because we have not included any baseline covariates, the estimator in `mediate()` is only consistent for the indirect effect if all confounders are included in the models in `mediate()`.
 - This estimator is in any case consistent for $\Delta - \Delta_S$ in the definition of PTE_{WT} when no baseline covariates have been included.
 - This estimator is in any case consistent for $\Delta - E\{\Delta_S(U)\}$ in the definition of PTE_{WT}^X when baseline covariates have been included.
- Average direct effect, ADE `(treated)`. As for the indirect effect, the estimator in `mediate()` is only consistent for the direct effect if all confounders have been included in the models in `mediate()`.
 - This estimator is in any case consistent for Δ_S in the definition of PTE_{WT} when no baseline covariates have been included.
 - This estimator is in any case consistent for $E\{\Delta_S(U)\}$ in the definition of PTE_{WT}^X when baseline covariates have been included.
- Total effect. This estimator is in any case consistent for Δ because treatment is randomized.
- Proportion mediated, Prop. Mediated `(treated)`. This estimator is the ratio of ACME `(treated)` and Total effect.
 - This is a consistent estimator for PTE_{WT} when no baseline covariates have been included.
 - This is a consistent estimator for PTE_{WT}^X when baseline covariates have been included.

```
PTE_WT = mediate(
  # Model for the mediator, i.e., the surrogate.
  model.m = surrogate_marg_model,
  # Model for the outcome, i.e., the true endpoint.
  model.y = true_endpoint_marg_model,
  # Number of MC samples for the numerical approximation.
  sims = 1e3,
  # Name of the treatment variable.
  treat = "Z",
  # Name of the mediator (surrogate) variable.
  mediator = "S")
```

```
)
summary(PTE_WT)
```

```
##
## Causal Mediation Analysis
##
## Quasi-Bayesian Confidence Intervals
##
##           Estimate 95% CI Lower 95% CI Upper p-value
## ACME (control)      0.0838      0.0788      0.09 <2e-16 ***
## ACME (treated)      0.1011      0.0929      0.11 <2e-16 ***
## ADE (control)       -0.0323     -0.0448     -0.02 <2e-16 ***
## ADE (treated)       -0.0150     -0.0206     -0.01 <2e-16 ***
## Total Effect        0.0688      0.0605      0.08 <2e-16 ***
## Prop. Mediated (control) 1.2176      1.1210      1.34 <2e-16 ***
## Prop. Mediated (treated) 1.4700      1.2605      1.72 <2e-16 ***
## ACME (average)      0.0924      0.0864      0.10 <2e-16 ***
## ADE (average)       -0.0237     -0.0326     -0.01 <2e-16 ***
## Prop. Mediated (average) 1.3438      1.1910      1.53 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Sample Size Used: 20000
##
##
## Simulations: 1000
```

```
PTE_WT_X = mediate(
  model.m = surrogate_model,
  model.y = true_endpoint_model,
  sims = 1e3,
  treat = "Z",
  mediator = "S"
)
summary(PTE_WT_X)
```

```
##
## Causal Mediation Analysis
##
## Quasi-Bayesian Confidence Intervals
##
##           Estimate 95% CI Lower 95% CI Upper p-value
## ACME (control)      0.06561      0.06077      0.07 <2e-16 ***
## ACME (treated)      0.06194      0.05429      0.07 <2e-16 ***
## ADE (control)       0.00903     -0.00376      0.02  0.16
## ADE (treated)       0.00536     -0.00218      0.01  0.16
## Total Effect        0.07097      0.06294      0.08 <2e-16 ***
## Prop. Mediated (control) 0.92432      0.83231      1.03 <2e-16 ***
## Prop. Mediated (treated) 0.87046      0.72259      1.06 <2e-16 ***
## ACME (average)      0.06378      0.05769      0.07 <2e-16 ***
## ADE (average)       0.00719     -0.00297      0.02  0.16
## Prop. Mediated (average) 0.89739      0.77676      1.05 <2e-16 ***
```



```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Sample Size Used: 20000
##
##
## Simulations: 1000
```

Interpretation of Estimates

The results are summarized in the following table.

```
eval_trial_tbl = eval_trial_tbl %>%
  mutate(S_tilde = predict(true_endpoint_model, newdata = eval_trial_tbl %>%
    mutate(Z = 1), type = "response"))
eval_trial_tbl %>%
  group_by(Z) %>%
  summarize(predicted_proportion = mean(S_tilde), proportion = mean(infection_free)) %>%
  pivot_longer(cols = 2:3) %>%
  pivot_wider(names_from = "Z", values_from = "value") %>%
  mutate(Delta = `1` - `0`)
```

```
## # A tibble: 2 x 4
##   name                '0'    '1' Delta
##   <chr>              <dbl> <dbl> <dbl>
## 1 predicted_proportion 0.871 0.932 0.0610
## 2 proportion          0.862 0.932 0.0700
```

Application Trials

Data Generation

Prediction of Treatment Effects