**Appendix 1: R code.**

# Activation of required libraries

library(rms)

library(mvtnorm)

library(arm)

library(rdd)

# Hypothetical dataset

n\_patients <- 100

R2 <- 0.0

corr <- sqrt(R2)

treatment\_effect <- -10

n\_sim <- 10000

Est\_Effect <- matrix(nrow = n\_sim, ncol=6)

Std\_error <- matrix(nrow = n\_sim, ncol=6)

DF <- matrix(nrow = n\_sim, ncol=6)

# Variance of prognosis (sigma1), outcome (sigma2)

sigma1 <- 4

sigma2 <- 100

#Covariantie prognosis and outcome

sigma12 <- corr \* sqrt(sigma1) \* sqrt(sigma2)

# Mean of prognosis and outcome

mu <- c(10, 90)

# Covariance Matrix

sigma <- matrix(c(sigma1, sigma12, sigma12, sigma2), nrow = 2, byrow = TRUE)

for(i in 1:n\_sim){

dataset <- rmvnorm(n\_patients, mean = mu, sigma = sigma)

dataset <- as.data.frame(dataset)

names(dataset) <- c("prognosis", "outcome")

## Randomized Controlled Trial, all patients randomized

# Treatment "Randomize all patients"

dataset$T\_RCT <- as.numeric(runif(n\_patients) <= 0.5)

# Outcome "Randomize all patients"

dataset$O\_RCT <- dataset$outcome

dataset$O\_RCT[dataset$T\_RCT == 1] <- dataset$outcome[dataset$T\_RCT == 1] + treatment\_effect

## Regression discontinuity, good prognosis control, poor prognosis treatment

#Treatment "Regression discontinuity design"

dataset$T\_RDC <- as.numeric(dataset$prognosis>10)

#Outcome "Regression discontinuity design"

dataset$O\_RDC <- dataset$outcome

dataset$O\_RDC[dataset$T\_RDC == 1] <- dataset$outcome[dataset$T\_RDC == 1] + treatment\_effect

fit\_RCT <- lm(O\_RCT ~ prognosis + T\_RCT, data = dataset)

fit\_RDC <- lm(O\_RDC ~ prognosis + T\_RDC, data = dataset)

fit\_RCT\_rcs <- ols(O\_RCT ~ rcs(prognosis) + T\_RCT, data = dataset, x=T, y=T)

fit\_RDC\_rcs <- ols(O\_RDC ~ rcs(prognosis) + T\_RDC, data = dataset, x=T, y=T)

fit\_RCT\_llr <- RDestimate(O\_RCT ~ prognosis, cutpoint = 10, data = dataset)

fit\_RDC\_llr <- RDestimate(O\_RDC ~ prognosis, cutpoint = 10, data = dataset)

Est\_Effect[i, ] <- c(fit\_RCT$coefficients[3], fit\_RDC$coefficients[3], fit\_RCT\_rcs$coefficients[6], fit\_RDC\_rcs$coefficients[6], fit\_RCT\_llr$est[1], fit\_RDC\_llr$est[1])

}

#Mean Effect estimate of treatment

colMeans(Est\_Effect)

#Mean squared error of effect estimate treatment

colMeans((Est\_Effect - treatment\_effect)^2)

**Appendix 2. Histogram of the assignment variable in both preDIVA and PROSPER in RCT data.**

1. **Histogram of mean systolic blood pressure (mmHg) at baseline in PreDIVA.**

****

1. **Histogram of age (years) at baseline in PreDIVA.**



**c) Histogram of total cholesterol level (mmol/L) at baseline in PROSPER.**

****

1. **Histogram of age (years) at baseline in PROSPER.**



**Table 5a. Relative efficiency of global treatment effect estimates in RD design in terms of required sample size (compared to global treatment effect estimate in RCT design) for different validation studies in preDIVA and PROSPER\*.**

 **preDIVA PROSPER**

|  |  |  |
| --- | --- | --- |
| RCT (linear adjustment) vs RD (RCS adjustment) | 6.25 1 | 9.0 1 |
| RCT (linear adjustment) vs RD (RCS adjustment) | 8.45 2 | 11.52 2 |

*\*Formula: (SE RD / SE RCT )2*

*1 Patient selection age ≤ 72 Tx- and BP > 72 Tx+*

*2 Patient selection age ≤ 74 Tx- and BP > 74 Tx+*

**Table 5b. Relative efficiency of local treatment effect estimates in RD design in terms of required sample size (compared to local treatment effect estimate in RCT design) for different validation studies in preDIVA and PROSPER\*.**

 **preDIVA PROSPER**

|  |  |  |
| --- | --- | --- |
| RCT (RCS adjustment) vs RD (local linear regression) | 3.56 1 | 1.04 3 |
| RCT (RCS adjustment) vs RD (local linear regression) | 3.78 2 | 0.72 4 |

*\*Formula: (SE RD / SE RCT )2*

*1 Patient selection BP ≤ 140 Tx- and BP > 140 Tx+*

*2 Patient selection BP ≤ 160 Tx- and BP > 160 Tx+*

*3 Patient selection cholesterol ≤ 5.0 Tx- and cholesterol > 5.0 Tx+*

*4 Patient selection cholesterol ≤ 5.5 Tx- and cholesterol > 5.5 Tx+*