T2D

2024-10-18

Differences in T2D diagnosed people

We look into the differences between the Treatment group and the placebo group in different scenarios. The effect of T2D on Death is in all simulations 1. For each setting we simulate where the effect of the drug on T2D is respectively 0, -0.5, -1.

We define a summary function

```
my_summary <- function(data) {
  # T2D events
  T2D_events <- data[Delta == 3]
  # T2D people
  T2D_peeps <- data[ID %in% T2D_events$ID]
  # Mean baseline L0
  L0Mean <- T2D_events[, .(L0_mean = mean(L0)), by = A0][order(A0)]
  # Mean Time of T2D diagnosis
  T_Mean <- T2D_events[, .(T_mean = mean(Time)), by = A0][order(A0)]
  # Number of T2D events in the two groups
  num_events <- T2D_events[,  .N, by = A0][order(A0)]

  # Setting T_0 to debut time of diabetes
  T2D_peeps[, Time_T2D := Time - min(Time), by = ID]
  # Removing the new Time 0
  T2D_peeps <- T2D_peeps[Delta != 3]

  # Proportion of treatment and placebo patients who have died before 1 year after T2D diagnose
  prop_treat <- nrow(T2D_peeps[Time_T2D < 1 & Delta == 1 & A0 == 1]) / length(unique(T2D_peeps[A0 == 1]$
  prop_plac <- nrow(T2D_peeps[Time_T2D < 1 & Delta == 1 & A0 == 0]) / length(unique(T2D_peeps[A0 == 0]$I

  table_output <- data.table("A0" = L0Mean$A0, "L0 mean" = L0Mean$L0_mean,
                             "T2D Time mean" = T_Mean$T_mean,
                             "Number of Events" = num_events$N,
                            "Prop dead 2 years after T2D" = c(prop_plac, prop_treat))


  return(table_output)
}
```

And histogram function

```
my_hist <- function(data0, data0.5, data1) {
  # T2D events
  T2D_events0 <- data0[Delta == 3]; T2D_events0.5 <- data0.5[Delta == 3]
  T2D_events1 <- data1[Delta == 3]

  my_hist <- gridExtra::grid.arrange(
```

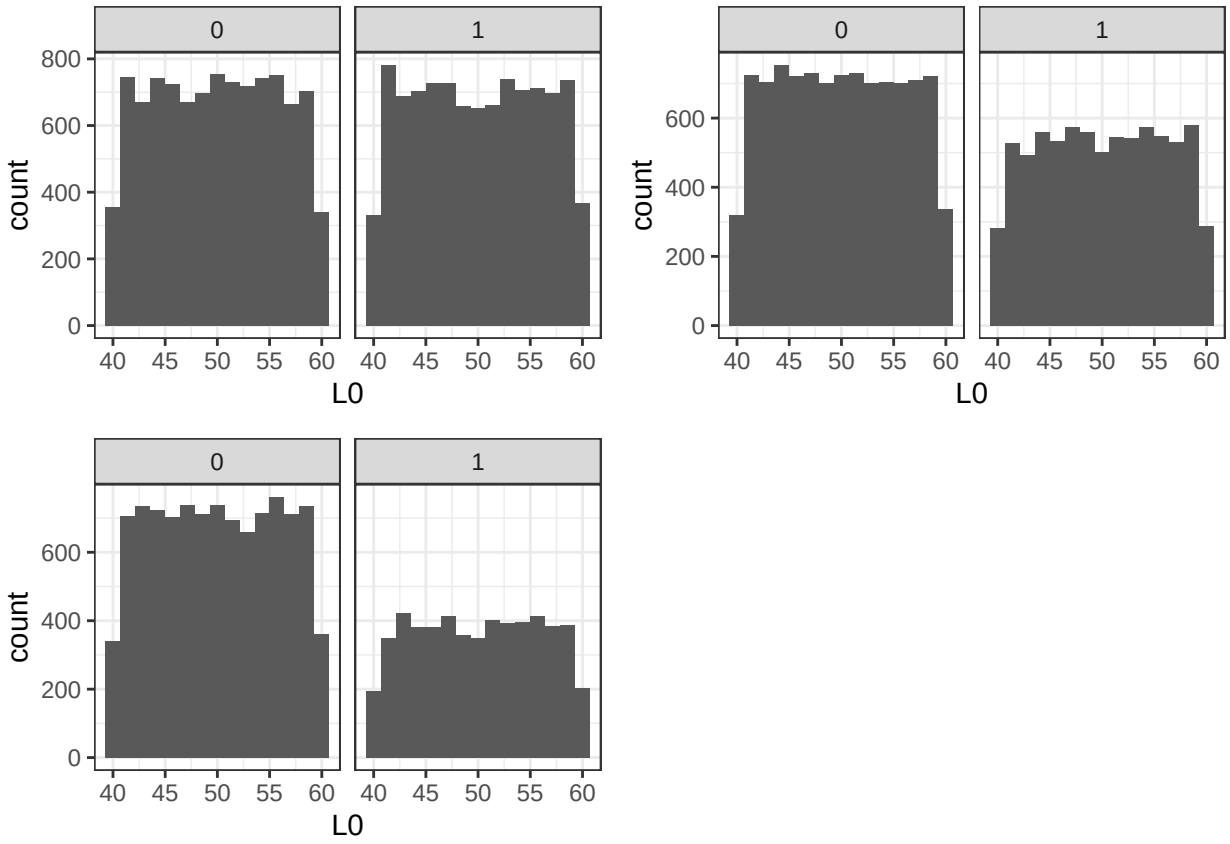
```
  ggplot(T2D_events0)+
    geom_histogram(aes(x = L0), bins = 15)+
    facet_grid(~A0),
  ggplot(T2D_events0.5)+
    geom_histogram(aes(x = L0), bins = 15)+
    facet_grid(~A0),
  ggplot(T2D_events1)+
    geom_histogram(aes(x = L0), bins = 15)+
    facet_grid(~A0),
  nrow = 2
)

  return(my_hist)
}
```

Scenario A: No effect of drug on death, L0 has no effect on neither T2D nor on Death.

```
N <- 4*10^4
data0_a <- sim_data_setting2(N = N, beta_L_D = 1, beta_A0_D = 0,
                             beta_L0_L = 0, beta_A0_L = 0,
                             cens = 0, beta_L0_D = 0)
data0.5_a <- sim_data_setting2(N = N, beta_L_D = 1, beta_A0_D = 0,
                             beta_L0_L = 0, beta_A0_L = -0.5,
                             cens = 0, beta_L0_D = 0)
data1_a <- sim_data_setting2(N = N, beta_L_D = 1, beta_A0_D = 0,
                             beta_L0_L = 0, beta_A0_L = -1, beta_L0_D = 0,
                             cens = 0)
```

First we look at the distribution of the covariate L_0:

```
my_hist(data0_a, data0.5_a, data1_a)
```




```
## TableGrob (2 x 2) "arrange": 3 grobs
##   z     cells    name           grob
## 1 1 (1-1,1-1) arrange gtable[layout]
## 2 2 (1-1,2-2) arrange gtable[layout]
## 3 3 (2-2,1-1) arrange gtable[layout]
```

Looks pretty evenly distributed.

```
my_summary(data0_a) |> kable()
```

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	49.97666	4.162244	10006	0.2788327
1	49.99871	4.191365	9883	0.2875645

```
my_summary(data0.5_a) |> kable()
```

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	49.97766	4.167739	9983	0.2890915
1	50.10102	5.099579	7644	0.2961800

```
my_summary(data1_a) |> kable()
```

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	50.02058	4.100676	10030	0.2866401
1	50.06785	5.932581	5429	0.2919506

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D

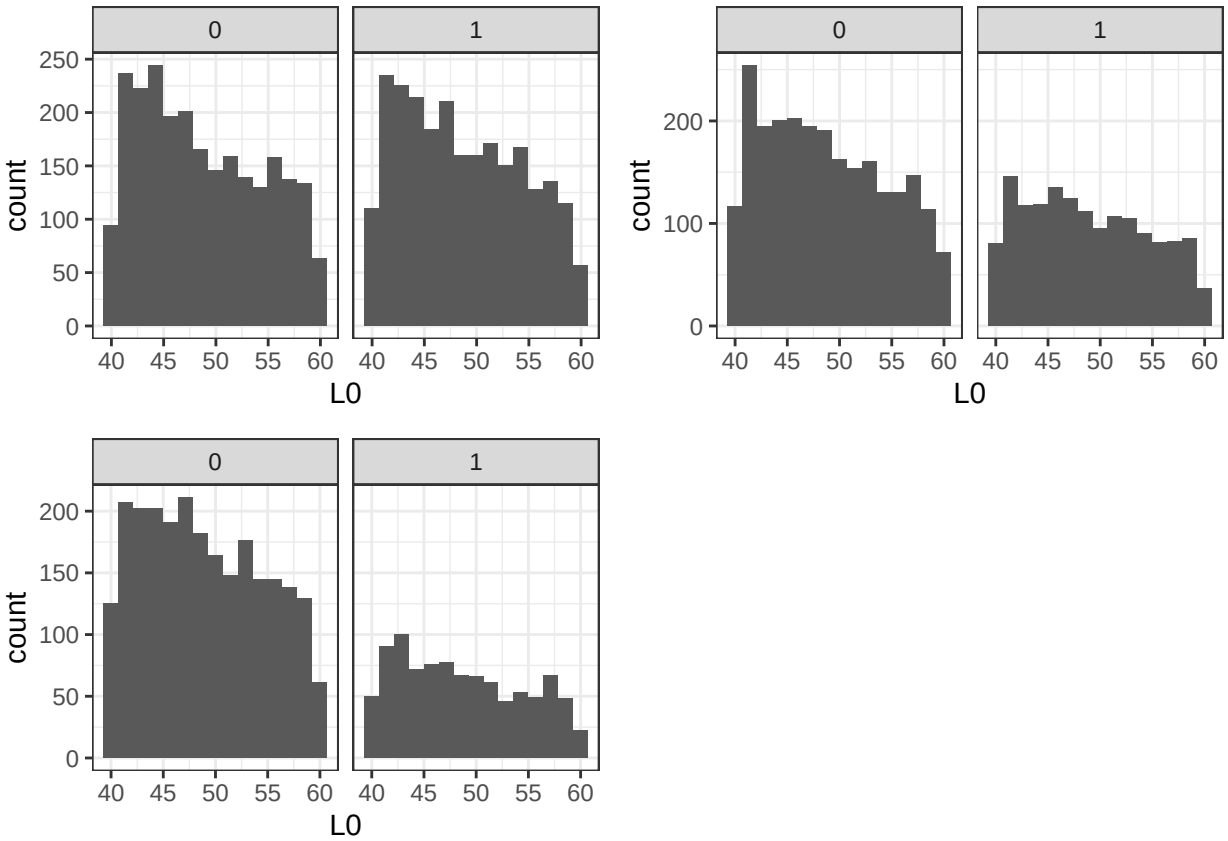
The L0 mean is roughly 50, the larger the effect of the drug, the later the T2D diagnose. The number of T2D diagnosed patients is smaller when the effect of there is an effect of the drug. The proportion of T2D patients dead two years after T2D diagnose is increasing slightly as the effect of the drug increases (in the placebo group a constant proportion 0.29 are dead).

Scenario B: No effect of drug on death, L0 has no effect on T2D but an (very large) effect on Death.

```
data0_b <- sim_data_setting2(N = N, beta_L_D = 1, beta_A0_D = 0,
                             beta_L0_L = 0, beta_A0_L = 0,
                             cens = 0, beta_L0_D = 2)
data0.5_b <- sim_data_setting2(N = N, beta_L_D = 1, beta_A0_D = 0,
                             beta_L0_L = 0, beta_A0_L = -0.5,
                             cens = 0, beta_L0_D = 2)
data1_b <- sim_data_setting2(N = N, beta_L_D = 1, beta_A0_D = 0,
                             beta_L0_L = 0, beta_A0_L = -1, beta_L0_D = 2,
                             cens = 0)
```

First we look at the distribution of the covariate L_0:

```
my_hist(data0_b, data0.5_b, data1_b)
```




```
## TableGrob (2 x 2) "arrange": 3 grobs
##   z     cells    name           grob
## 1 1 (1-1,1-1) arrange gtable[layout]
```

```
## 2 2 (1-1,2-2) arrange gtable[layout]
## 3 3 (2-2,1-1) arrange gtable[layout]
```

Now the distribution is uneven, with the T2D patients having a smaller $L0$. This might be due to the people with high L0 values being dead. But the effect is the same for placebo and treatment patients.

```
my_summary(data0_b) |> kable()
```

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	48.81499	1.190055	2426	0.8812861
1	48.72921	1.191384	2421	0.8773234

```
my_summary(data0.5_b) |> kable()
```

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	48.80511	1.186184	2423	0.8778374
1	48.86733	1.254049	1516	0.8970976

```
my_summary(data1_b) |> kable()
```

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	48.98570	1.214750	2426	0.8742786
1	48.69965	1.308529	944	0.8750000

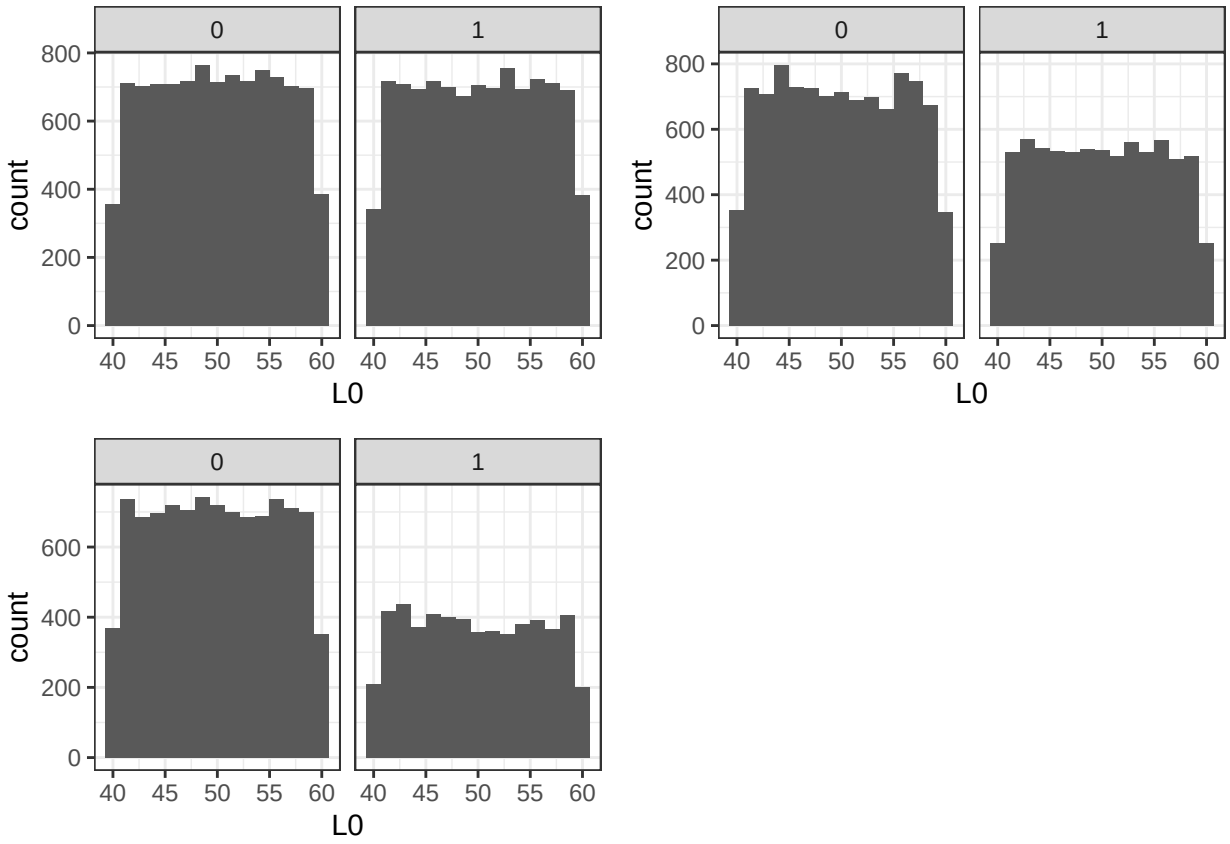
The mean of $L0$ is below 50, but not by much. The T2D time is again larger for the treatment group. The number of T2D events is decreasing for the T2D group as before. Prop dead 2 years after T2D is roughly the same in the two groups.

Scenario C: No effect of drug on death, L0 has an effect on T2D and also an effect on Death (of equal size).

```
data0_c <- sim_data_setting2(N = N, beta_L_D = 1, beta_A0_D = 0,
                             beta_L0_L = 1, beta_A0_L = 0,
                             cens = 0, beta_L0_D = 1)
data0.5_c <- sim_data_setting2(N = N, beta_L_D = 1, beta_A0_D = 0,
                             beta_L0_L = 1, beta_A0_L = -0.5,
                             cens = 0, beta_L0_D = 1)
data1_c <- sim_data_setting2(N = N, beta_L_D = 1, beta_A0_D = 0,
                             beta_L0_L = 1, beta_A0_L = -1, beta_L0_D = 1,
                             cens = 0)
```

First we look at the distribution of the covariate L_0:

```
my_hist(data0_c, data0.5_c, data1_c)
```




```
## TableGrob (2 x 2) "arrange": 3 grobs
##   z     cells    name           grob
## 1 1 (1-1,1-1) arrange gtable[layout]
## 2 2 (1-1,2-2) arrange gtable[layout]
## 3 3 (2-2,1-1) arrange gtable[layout]
```

Now the distribution seems to be even. The two effects are cancelling!

```
my_summary(data0_c) |> kable()
```

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	50.03995	1.676860	10107	0.5702978
1	50.04461	1.672313	9911	0.5661386

```
my_summary(data0.5_c) |> kable()
```

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	49.92212	1.684509	10049	0.5832421
1	49.94950	2.061311	7493	0.5800080

```
my_summary(data1_c) |> kable()
```

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	49.96240	1.663080	9950	0.5699497
1	49.82567	2.335037	5456	0.5714809

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D

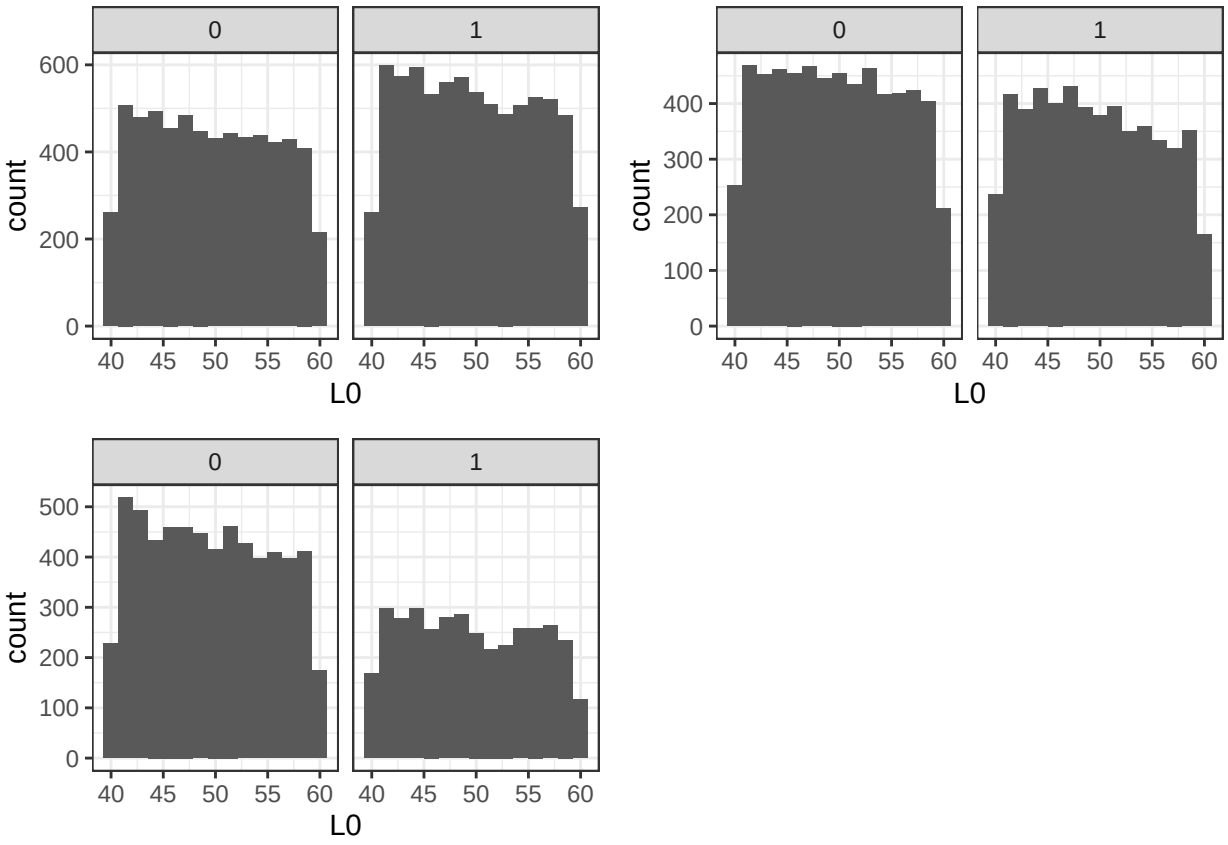
The same results as we found in scenario A.

Scenario D: Small effect of drug on Death (-0.3), L0 has a medium effect on T2D (0.7) and and a large effect on Death (1.5).

```
data0_c <- sim_data_setting2(N = N, beta_L_D = 1, beta_A0_D = -0.3,
                             beta_L0_L = 0.7, beta_A0_L = 0,
                             cens = 0, beta_L0_D = 1.5)
data0.5_c <- sim_data_setting2(N = N, beta_L_D = 1, beta_A0_D = -0.3,
                             beta_L0_L = 0.7, beta_A0_L = -0.5,
                             cens = 0, beta_L0_D = 1.5)
data1_c <- sim_data_setting2(N = N, beta_L_D = 1, beta_A0_D = -0.3,
                             beta_L0_L = 0.7, beta_A0_L = -1, beta_L0_D = 1.5,
                             cens = 0)
```

First we look at the distribution of the covariate L_0:

```
my_hist(data0_c, data0.5_c, data1_c)
```




```
## TableGrob (2 x 2) "arrange": 3 grobs
##   z     cells    name           grob
## 1 1 (1-1,1-1) arrange gtable[layout]
## 2 2 (1-1,2-2) arrange gtable[layout]
## 3 3 (2-2,1-1) arrange gtable[layout]
```

Since the effects are not symmetric, we see again find an uneven distribution, allbeit a less extreme one then the one from scenario B.

```
my_summary(data0_c) |> kable()
```

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	49.63329	1.463734	6345	0.7453113
1	49.71740	1.723967	7528	0.6478480

```
my_summary(data0.5_c) |> kable()
```

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	49.74067	1.422734	6227	0.7454633
1	49.49337	2.017159	5344	0.6519461

```
my_summary(data1_c) |> kable()
```

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	49.56742	1.461262	6134	0.7360613
1	49.56362	2.208004	3684	0.6463084

Pretty much the same results as in B. But this time with the proportion of dead after T2D diagnose smaller for the treatment group. But with the proportion of dead incresing as the effect of the drug increases. This must be due to the later T2D diagnose (?).

Larger effect of time on intensity of death

We now investigate how the estimates change if we increase the time effect on the intensity. Remember that we used the Weibull hazard as the baseline hazard:

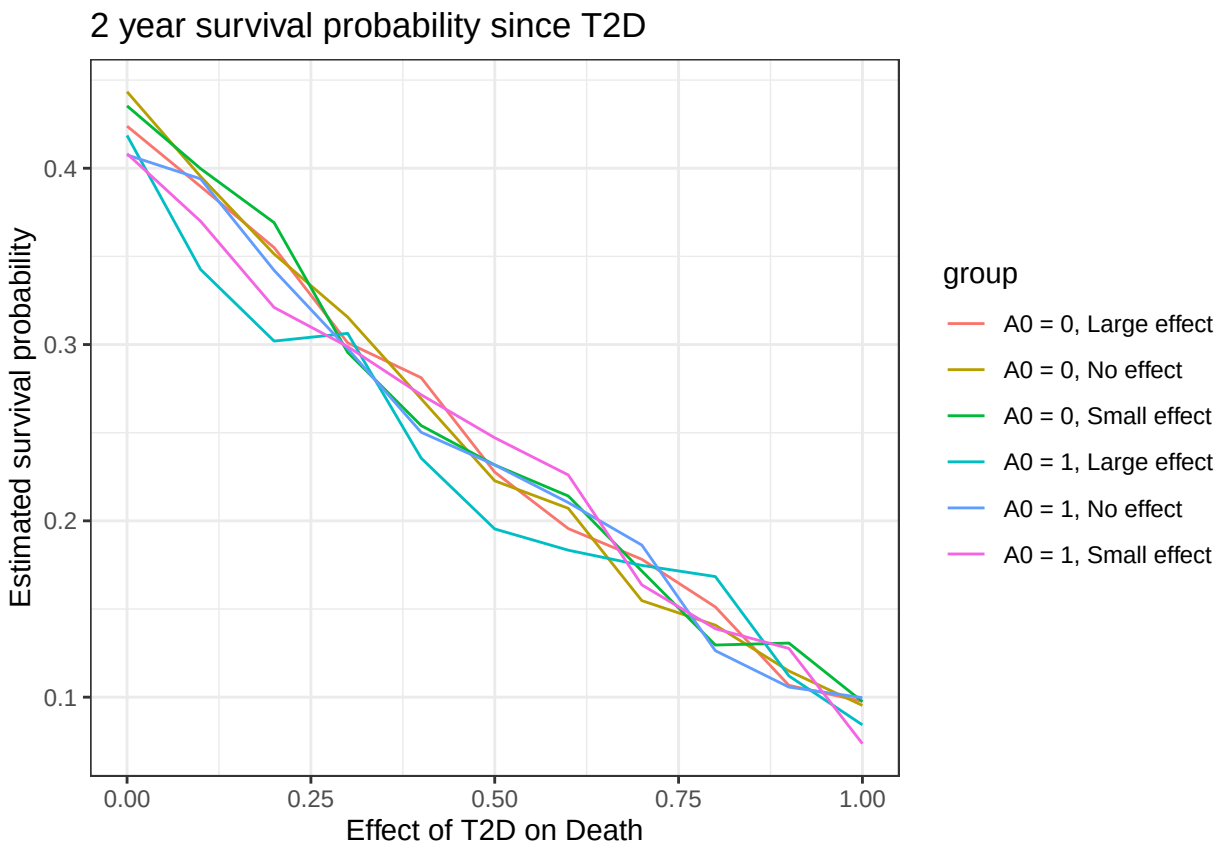
$$\alpha(t) = \eta\nu t^{nu-1}$$

We now increase ν from 1.1 to 1.3.

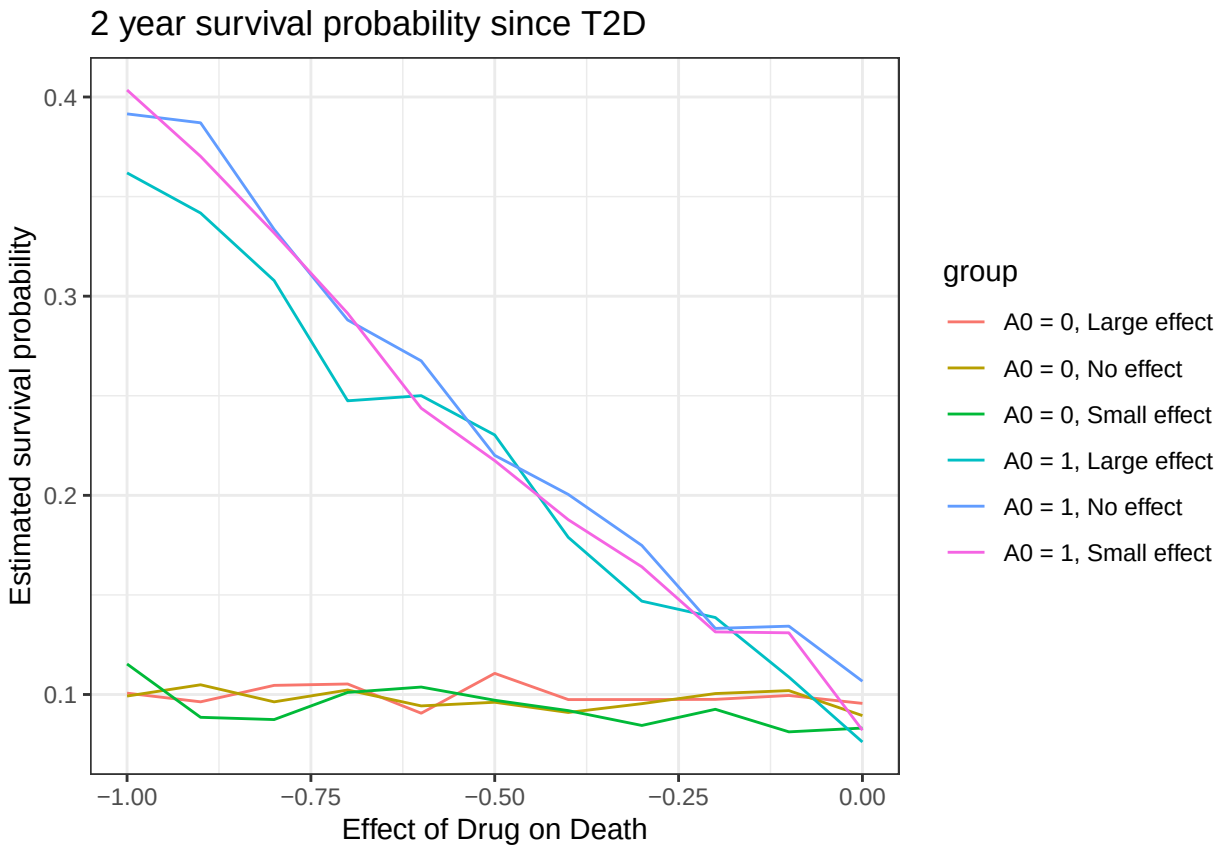
```
N <- 8000

estimator2 <- function(data, N) {
    # Finding all the T2D people
    T2D_events <- data[Delta == 3]
    T2D_peeps <- data[ID %in% T2D_events$ID]

    # Setting T_0 to debut time of diabetes
    T2D_peeps[, Time_T2D := Time - min(Time), by = ID]

    # Removing the new Time 0
    T2D_peeps <- T2D_peeps[Delta != 3]

    # Creating a status variable
    T2D_peeps[, Status := Delta == 1]

    # Kaplan meyer fit
```

```
    fit <- prodlim(Hist(Time_T2D, Status) ~ A0, data = T2D_peeps)

    # Save estimate of survival probability
    preds <- predict(fit, times = 2, newdata = data.frame(A0 = c(0,1)))
    return(c(preds$`A0=0`, preds$`A0=1`))
}

res5 <- compare_effects(estimator = estimator2, N = N, beta_L_D = seq(0,1,by = 0.1), nu = rep(1.3,4))
res6 <- compare_effects(estimator = estimator2, N = N, beta_A0_D = seq(-1,0,by = 0.1), nu = rep(1.3,4))
```

```
plot_compare(res5, diff_betas= seq(0,1,by = 0.1))+
      ylab("Estimated survival probability")+
      xlab("Effect of T2D on Death")+
      labs(title = "2 year survival probability since T2D")
```




```
plot_compare(res6, diff_betas= seq(-1,0,by = 0.1))+
      ylab("Estimated survival probability")+
      xlab("Effect of Drug on Death")+
      labs(title = "2 year survival probability since T2D")
```



Might there be a slight difference between the three curves? With the estimated survival probability being smallest when the drug has larger effect?