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(
```

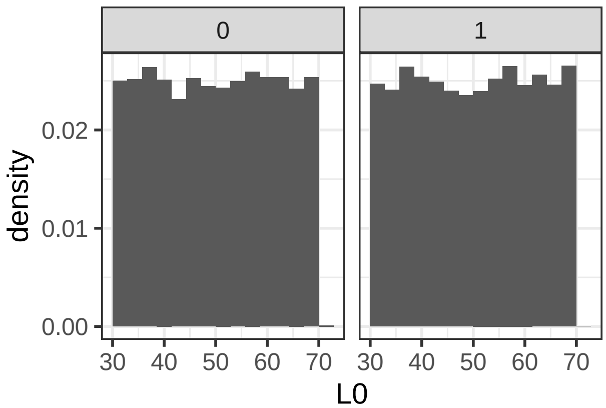
```r
  ggplot(T2D_events0)+
    geom_histogram(aes(x = L0, y =..density..), bins = 15)+
    facet_grid(~A0),
  ggplot(T2D_events0.5)+
    geom_histogram(aes(x = L0, y =..density..), bins = 15)+
    facet_grid(~A0),
  ggplot(T2D_events1)+
    geom_histogram(aes(x = L0, y =..density..), bins = 15)+
  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

  return(my_hist)
}
```

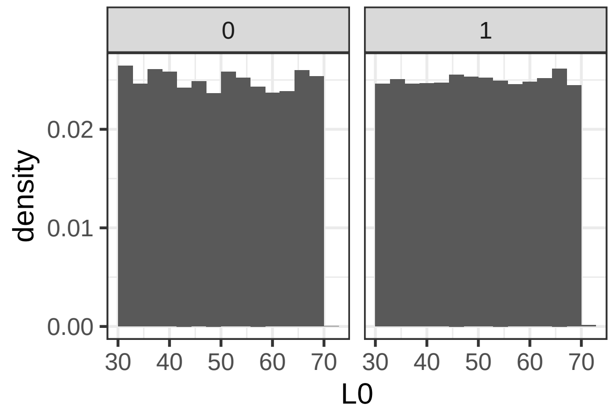
Scenario A: No effect of drug on death, L0 has no effect on neither T2D nor on Death.

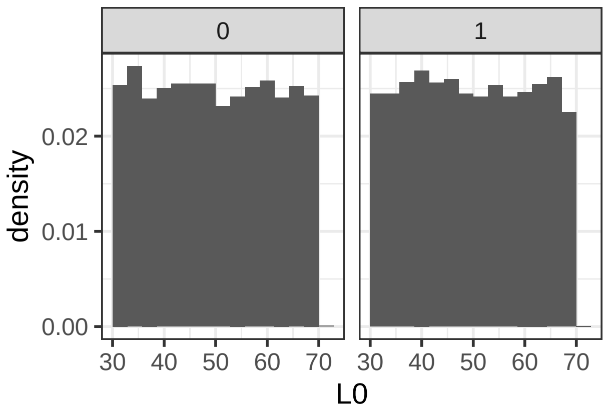
```r
N <- 4*10^4
data0_a <- sim_data_setting2(N = N, beta_L_D = 1, beta_A0_D = 0,
                             beta_L0_L = 0, beta_A0_L = 0,
                             cens = 0, beta_L0_D = 0)
data0.5_a <- sim_data_setting2(N = N, beta_L_D = 1, beta_A0_D = 0,
                             beta_L0_L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

```
## Warning: The dot-dot notation (`..density..`) was deprecated in ggplot2 3.4.0.
## i Please use `after_stat(density)` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```







Looks pretty evenly distributed.

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

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	49.99210	4.170998	9892	0.2910433
1	50.10097	4.176898	9935	0.2838450

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

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	49.88151	4.192733	10100	0.2956436
1	50.03400	5.072929	7494	0.2953029

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

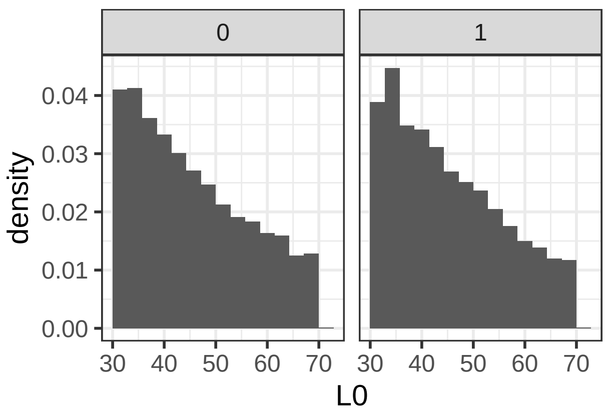
A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	49.84838	4.165032	9792	0.2831904
1	49.85908	5.822444	5454	0.2975798

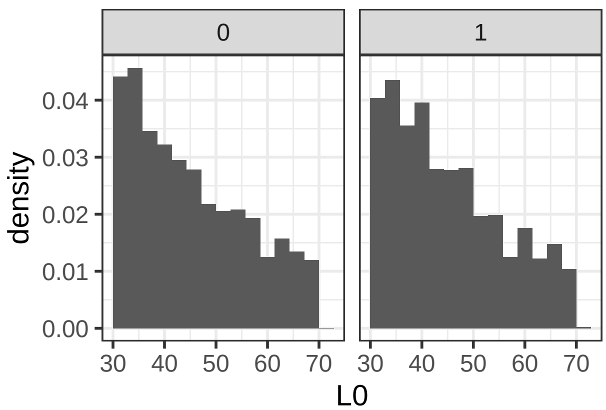
The L0 mean is roughly 50, the larger the effect of the drug, the later the T2D diagnose. The number of T2D diagnosed patients is smaller when the effect of there is an effect of the drug. The proportion of T2D patients dead two years after T2D diagnose is increasing slightly as the effect of the drug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

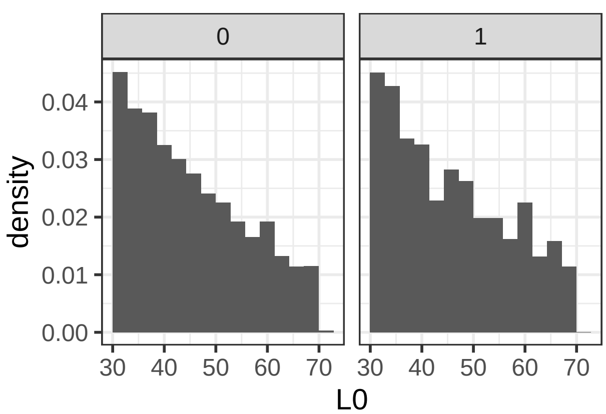
Scenario B: No effect of drug on death, L0 has no effect on T2D but an (very large) effect on Death.

```r
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:







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.

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

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	45.60139	1.346031	2637	0.8339022
1	45.36836	1.360712	2575	0.8310680

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

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	45.26060	1.416443	2553	0.8319624
1	45.18251	1.438432	1655	0.8241692

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

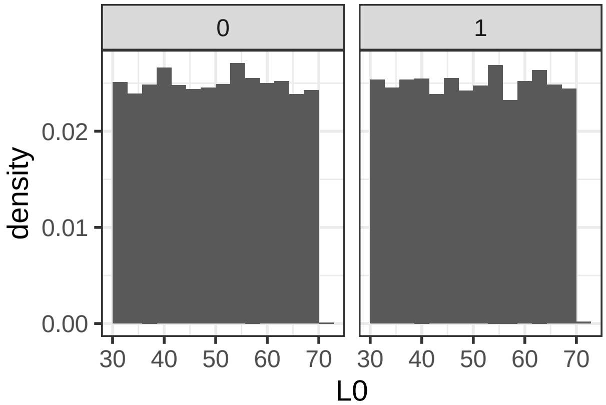
A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	45.30414	1.390578	2518	0.8264496
1	45.63330	1.486320	1041	0.8424592

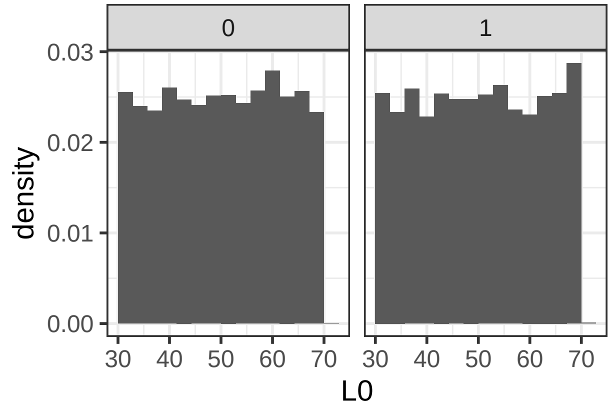
The mean of $L0$ is below 50, but not by much. The T2D time is again larger for the treatment group. The number of T2D events is decreasing for the T2D group as before. Prop dead 2 years after T2D is roughly the same in the two groups.

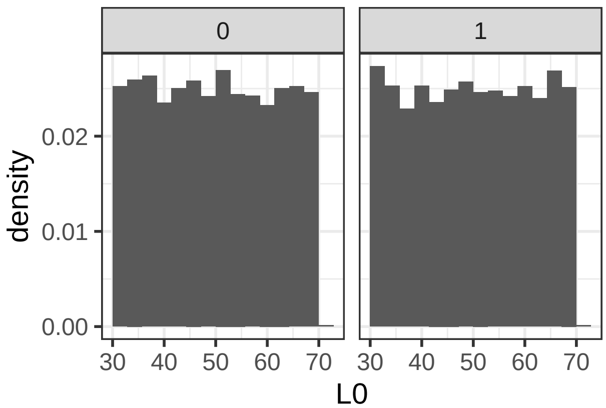
Scenario C: No effect of drug on death, L0 has an effect on T2D and also an effect on Death (of equal size).

```r
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:







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	49.97241	1.734152	9981	0.5826070
1	49.98801	1.718800	10013	0.5835414

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

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	50.08671	1.709456	9893	0.5685838
1	50.23829	2.079508	7541	0.5848031

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

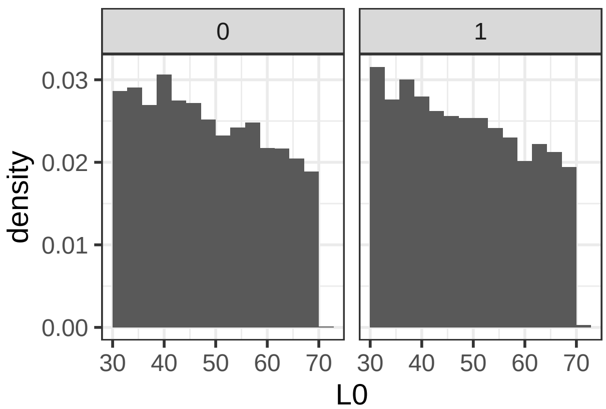
A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	49.83330	1.748779	9819	0.5699155
1	49.99116	2.396255	5312	0.5717244

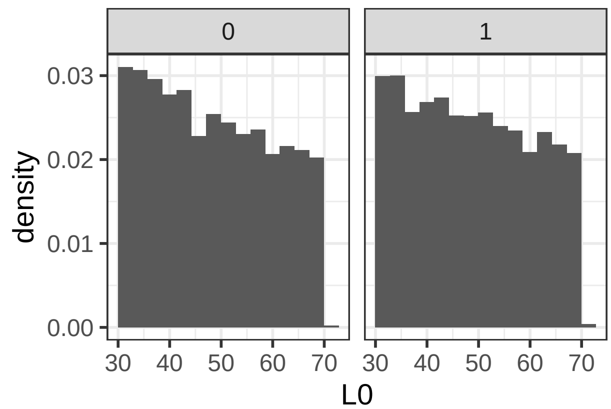
The same results as we found in scenario A.

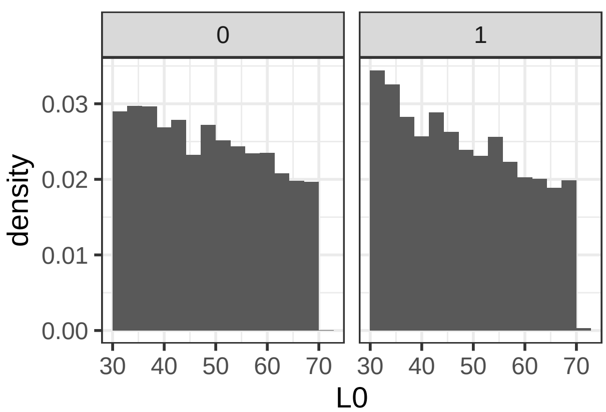
Scenario D: Small effect of drug on Death (-0.3), L0 has a medium effect on T2D (0.7) and and a large effect on Death (1.5).

```
data0_d <- 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_d <- 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_d <- 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:







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_d) |> kable()
```

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	48.54557	1.501691	6248	0.7299936
1	48.47991	1.821216	7436	0.6308499

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

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	48.37687	1.509907	6243	0.7195259
1	48.77836	2.090380	5395	0.6474513

```
my_summary(data1_d) |> kable()
```

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	48.52905	1.537313	6278	0.7328767
1	48.02709	2.376790	3531	0.6304163

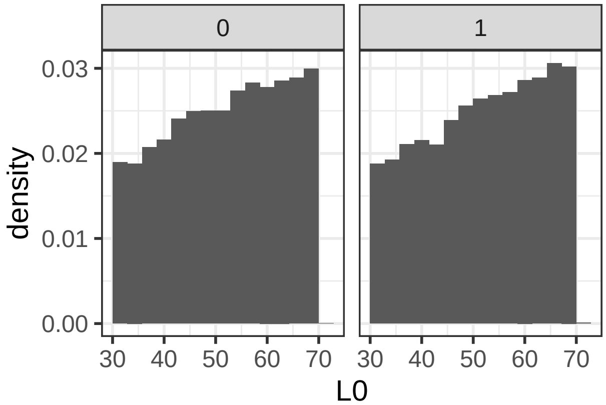
Pretty much the same results as in B. But this time with the proportion of dead after T2D diagnose smaller for the treatment group. But with the proportion of dead incresing as the effect of the drug increases. This must be due to the later T2D diagnose (?).

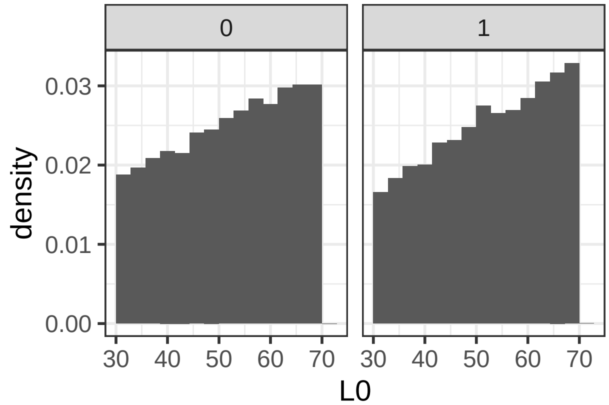
Scenarie E: L0 har en effekt KUN på T2D (2), No effect of drug on death, Effect of drug on T2D: 0, -0.5, -1.

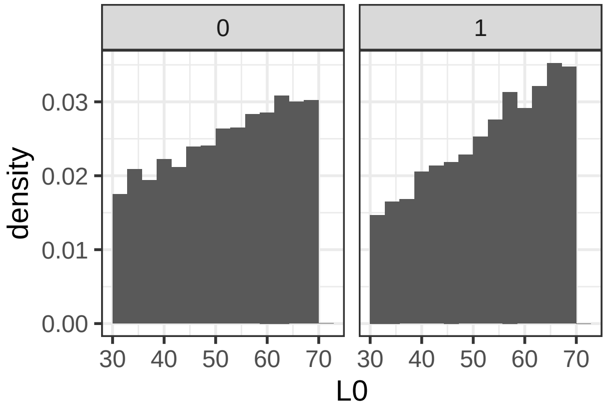
We run the simulation so that the intensity of death is largely effected by death. We increase the Weibull parameters of the death intensity from 0.1 to 0.3 and from 1.1 to 1.3.

```
data0_e <- sim_data_setting2(N = N, beta_L_D = 1, beta_A0_D = 0,
                           beta_L0_L = 2, beta_A0_L = 0,
                           cens = 0, beta_L0_D = 0,
                           eta = c(0.1,0.3,0.1,0.1), nu = c(1.1,1.3,1.1,1.1))
data0.5_e <- sim_data_setting2(N = N, beta_L_D = 1, beta_A0_D = 0,
                           beta_L0_L = 2, beta_A0_L = -0.5,
                           cens = 0, beta_L0_D = 0,
                           eta = c(0.1,0.3,0.1,0.1), nu = c(1.1,1.3,1.1,1.1))
data1_e <- sim_data_setting2(N = N, beta_L_D = 1, beta_A0_D = 0,
                           beta_L0_L = 2, beta_A0_L = -1, beta_L0_D = 0,
                           cens = 0,
                           eta = c(0.1,0.3,0.1,0.1), nu = c(1.1,1.3,1.1,1.1))
```

First we look at the distribution of the covariate L_0:







The distributions looks uneven. Increasingly as the effect of $A0$ on L increases.

```
my_summary(data0_e) |> kable()
```

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	51.63445	0.8429895	13735	0.6674190
1	51.77740	0.8537621	13794	0.6660867

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

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	51.75852	0.8437268	13676	0.6655455
1	52.25418	1.0693643	11312	0.6835219

```
my_summary(data1_e) |> kable()
```

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	51.89309	0.8465239	13810	0.6606807
1	53.00312	1.2935280	8794	0.6961565

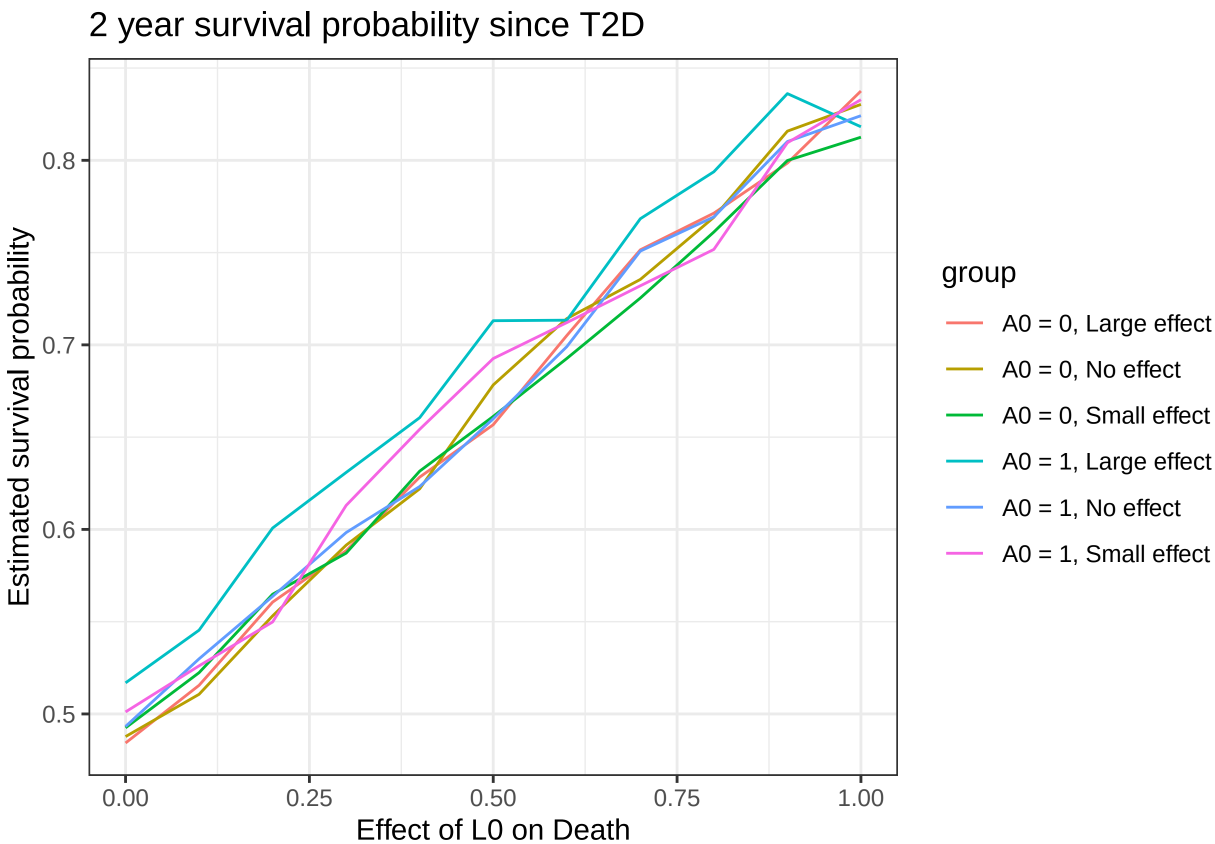
The L0 mean appears to be increasing for the Treatment group. For Treatment patients the Time of T2D increases substantially and the number of T2D patients decreases as the effect of the drug increases. The proportion of dead in the treatment group increases as the effect of the drug increases.

Kaplan-Meier estimates of survival for scenario E

We let the effect of $L0$ on death vary, and estimate the survival probability with KM.

```r
N <- 8000

estimator2 <- function(data, N) {
  # T2D events
  T2D_events <- data[Delta == 3]
  # T2D people
  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]

  # 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]$

    return(c(prop_plac, prop_treat))
}

res5 <- compare_effects(estimator = estimator2, N = N, beta_L_D = 0.5,
                        eta = c(0.1,0.3,0.1,0.1), nu = c(1.1,1.3,1.1,1.1),
                        beta_L0_L = 2, beta_L0_D = seq(0,1,by = 0.1), beta_A0_D = 0)
```

```r
plot_compare(res5, diff_betas= seq(0,1,by = 0.1))+
      ylab("Estimated survival probability")+
      xlab("Effect of L0 on Death")+
      labs(title = "2 year survival probability since T2D")
```



There is a slight difference between the curves.