

# 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]$ID))  
  prop_plac <- nrow(T2D_peeps[Time_T2D < 1 & Delta == 1 & A0 == 0]) / length(unique(T2D_peeps[A0 == 0]$ID))  
  
  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(  
    # Histograms for each group  
    # ...  
  )  
}
```

```

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,  $L_0$  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_LO_L = 0, beta_A0_L = 0,
                             cens = 0, beta_LO_D = 0)
data0.5_a <- sim_data_setting2(N = N, beta_L_D = 1, beta_A0_D = 0,
                               beta_LO_L = 0, beta_A0_L = -0.5,
                               cens = 0, beta_LO_D = 0)
data1_a <- sim_data_setting2(N = N, beta_L_D = 1, beta_A0_D = 0,
                             beta_LO_L = 0, beta_A0_L = -1, beta_LO_D = 0,
                             cens = 0)

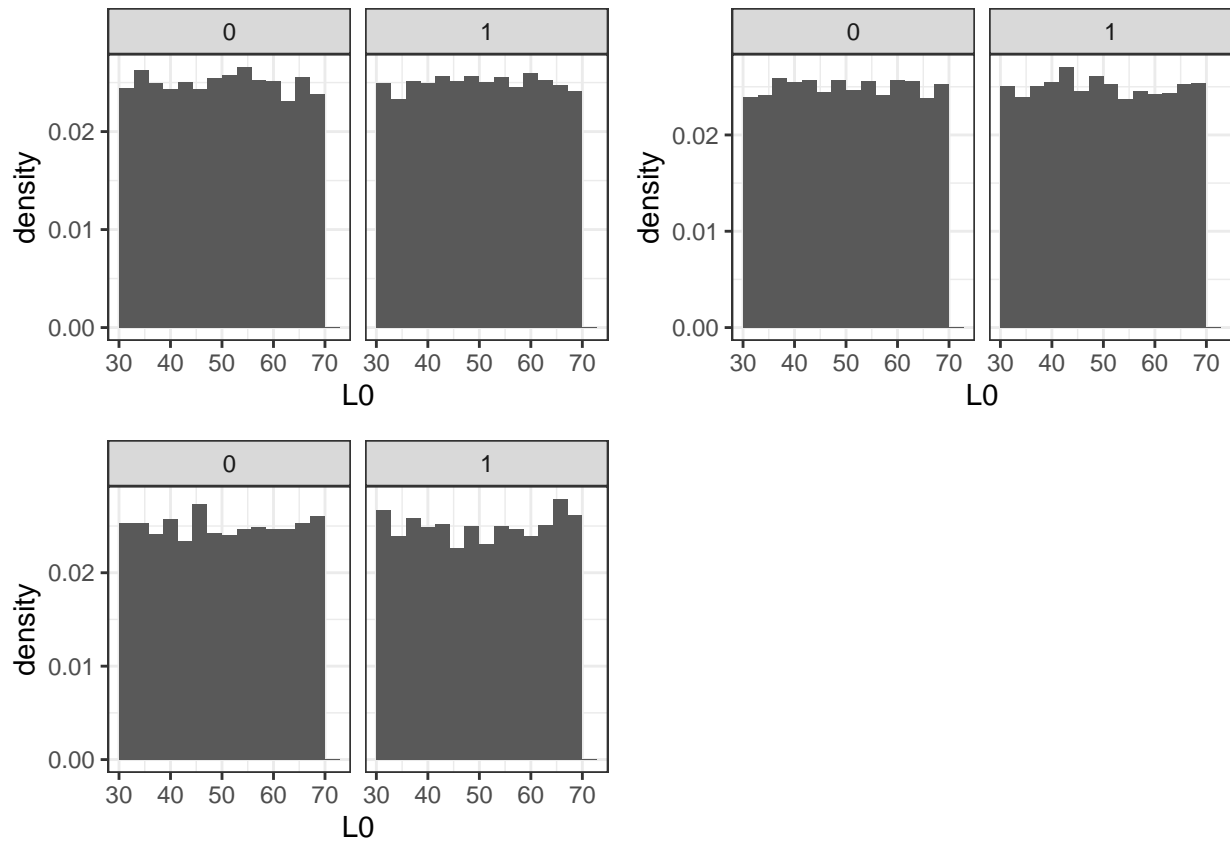
```

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.

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

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	49.93441	4.168585	10222	0.2967130
1	50.04394	4.206237	9852	0.2903979

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

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	50.03919	4.174133	10104	0.2832542
1	49.94873	5.241034	7471	0.2917949

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

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	50.02495	4.199072	9898	0.2851081
1	50.10374	5.920563	5350	0.3013084

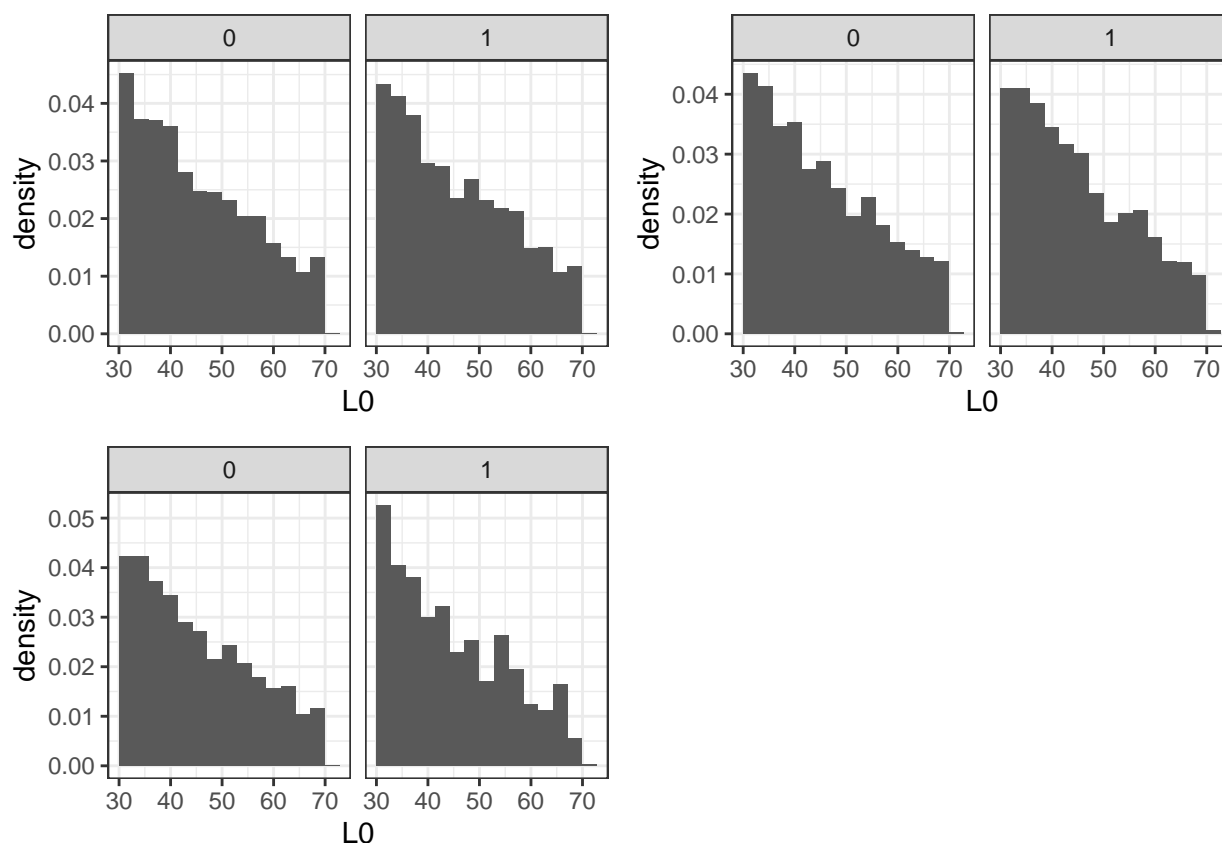
The L0 mean is roughly 50, the larger the effect of the drug, the later the T2D diagnose. The number of T2D diagnosed patients is smaller when the effect of there is an effect of the drug. The proportion of T2D patients dead two years after T2D diagnose is increasing slightly as the effect of the drug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

placebo group a constant proportion 0.29 are dead).

**Scenario B: No effect of drug on death,  $L_0$  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  $L_0$ :



Now the distribution is uneven, with the T2D patients having a smaller  $L_0$ . This might be due to the people with high  $L_0$  values being dead. But the effect is the same for placebo and treatment patients.

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

A0	$L_0$ mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	45.42946	1.336724	2571	0.8101906
1	45.45741	1.358705	2534	0.8188635

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

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	45.37744	1.344828	2587	0.8361036
1	45.15131	1.492045	1614	0.8246592

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

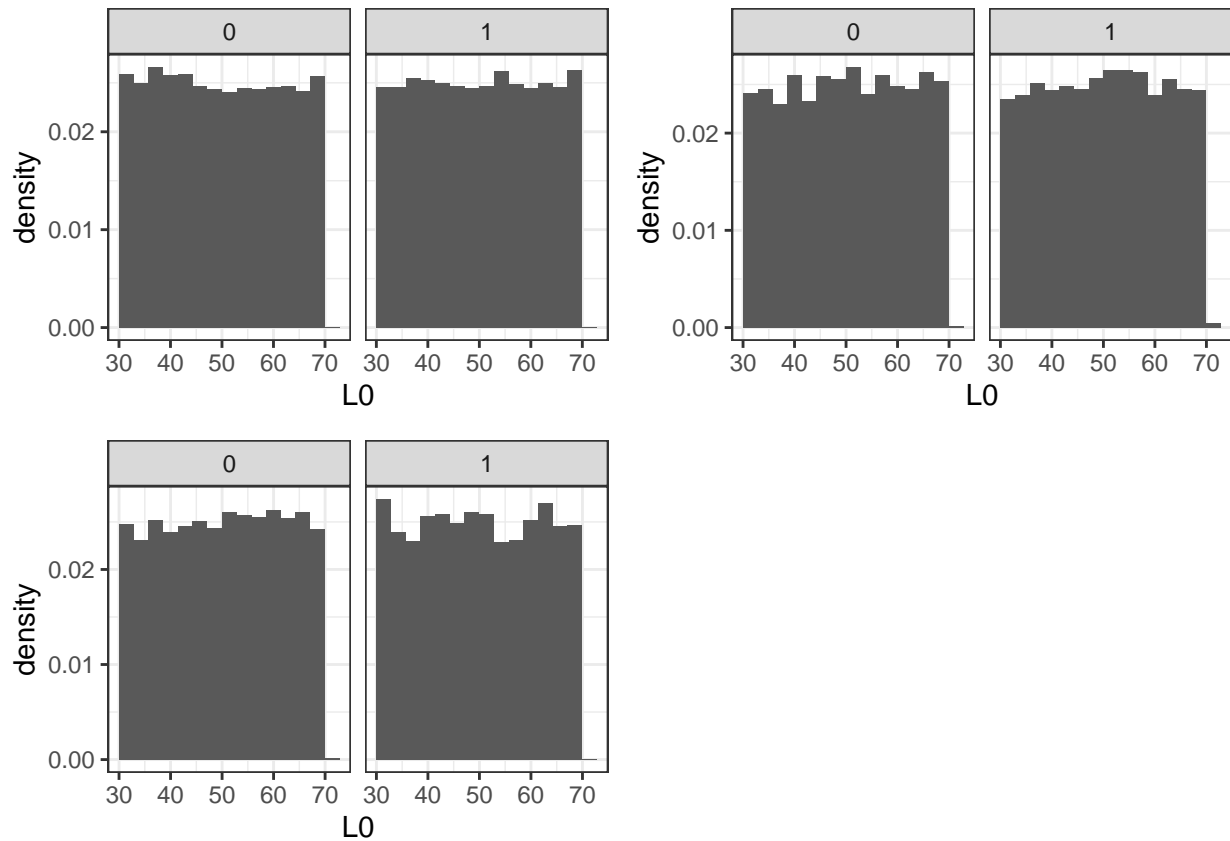
A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	45.26834	1.384145	2523	0.8315497
1	44.63899	1.554268	1066	0.8424015

The mean of  $L_0$  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,  $L_0$  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  $L_0$ :



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	49.80151	1.709435	10160	0.5715551
1	50.06721	1.718651	9983	0.5812882

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

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	50.20221	1.724887	10078	0.5777932
1	50.14693	2.034094	7507	0.5827894

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

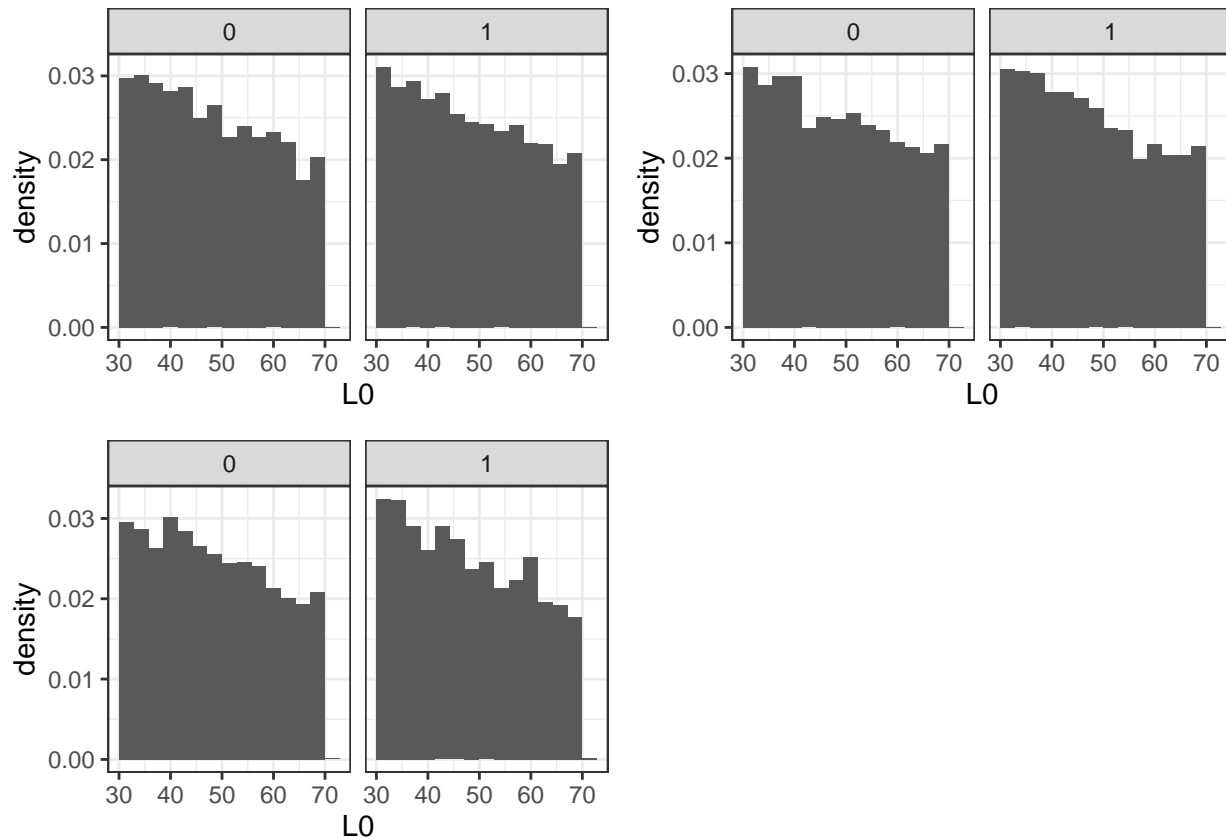
A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	50.18306	1.701661	9911	0.5713853
1	49.92600	2.422109	5407	0.5881265

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, albeit a less extreme one than 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	48.37737	1.499445	6297	0.7233603
1	48.51269	1.809247	7326	0.6255801

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

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	48.61404	1.539190	6313	0.7330904
1	48.32772	2.107136	5383	0.6375627

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

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	48.51262	1.537927	6246	0.7340698
1	48.03853	2.350183	3682	0.6328083

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 increasing as the effect of the drug increases. This must be due to the later T2D diagnose (?).

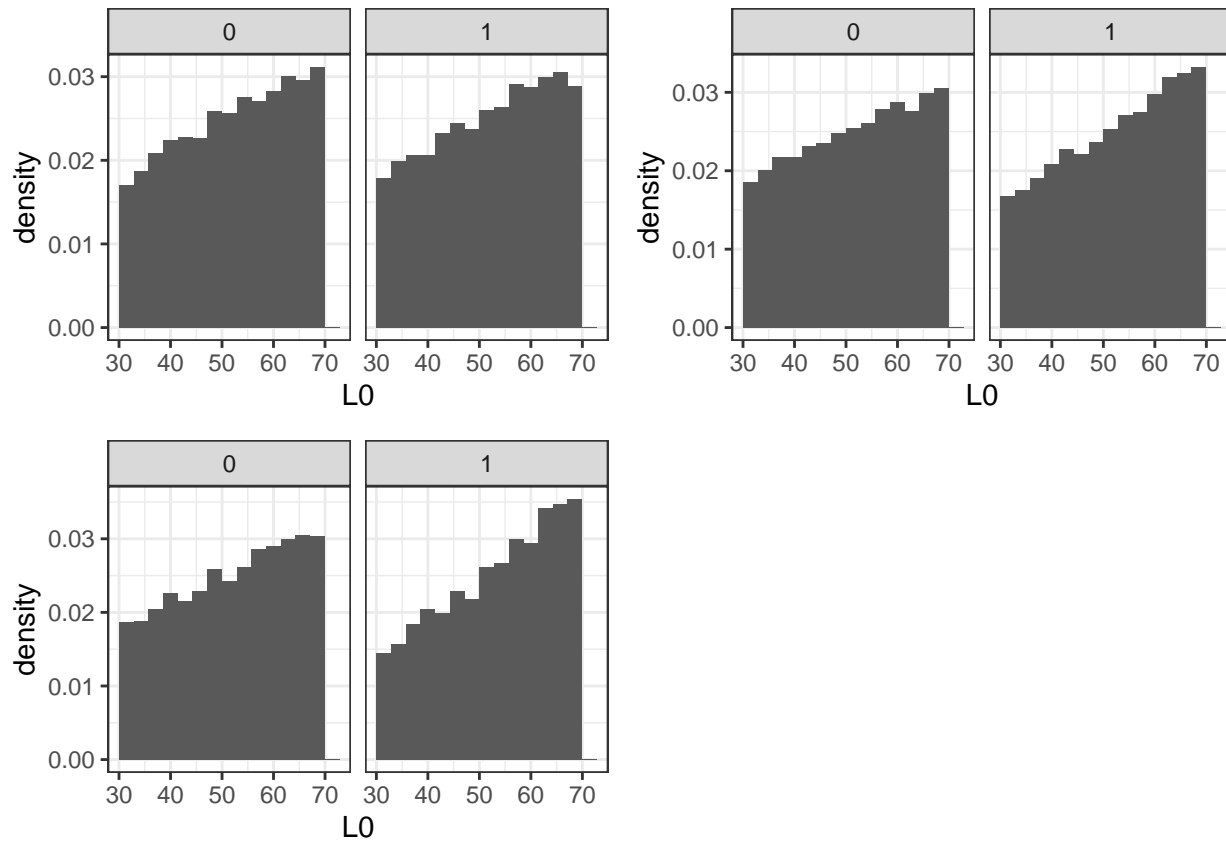
### Scenarie E: L0 har en effekt KUN på T2D (2), No effect of drug on death, Effect of drug on T2D: 0, -0.5, -1.

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  $A_0$  on  $L$  increases.

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

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	51.88826	0.8465128	13761	0.6647773
1	51.79323	0.8505583	13587	0.6674763

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

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	51.61983	0.853309	13799	0.6641786
1	52.45660	1.089239	11204	0.6870760

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

A0	L0 mean	T2D Time mean	Number of Events	Prop dead 2 years after T2D
0	51.85686	0.8410364	13730	0.6587036
1	53.06807	1.2916561	8677	0.7012792

The  $L_0$  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  $L_0$  on death vary, and estimate the survival probability with KM.

```
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]

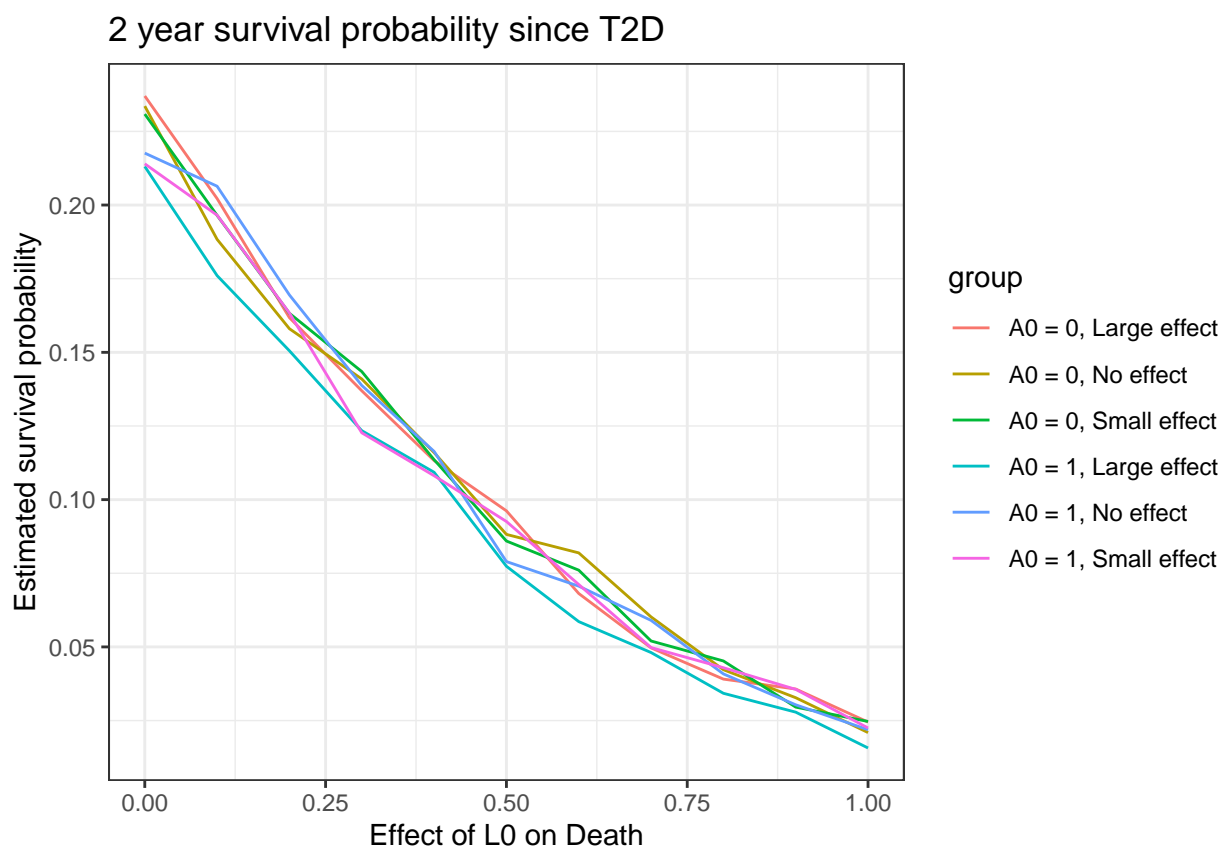
  # Creating a status variable
  T2D_peeps[, Status := Delta == 1]

  # 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 = 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)

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.