BIOST537_Project

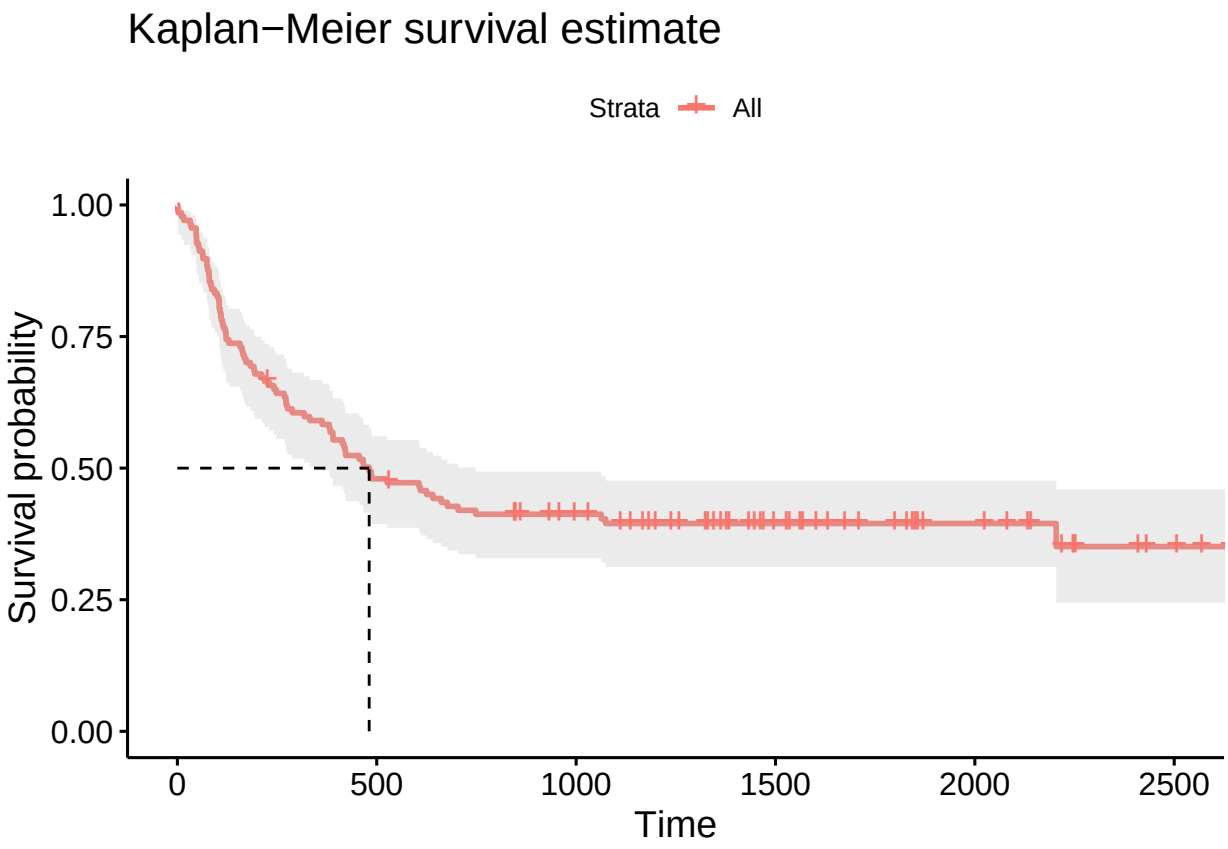
Dante Ramirez, Nada Ali, Jiayu Sun, Machi Kaneko

2023-03-13

Directive 1

```
## [1] 782.0292
```

```
## [1] 0.3941606
```

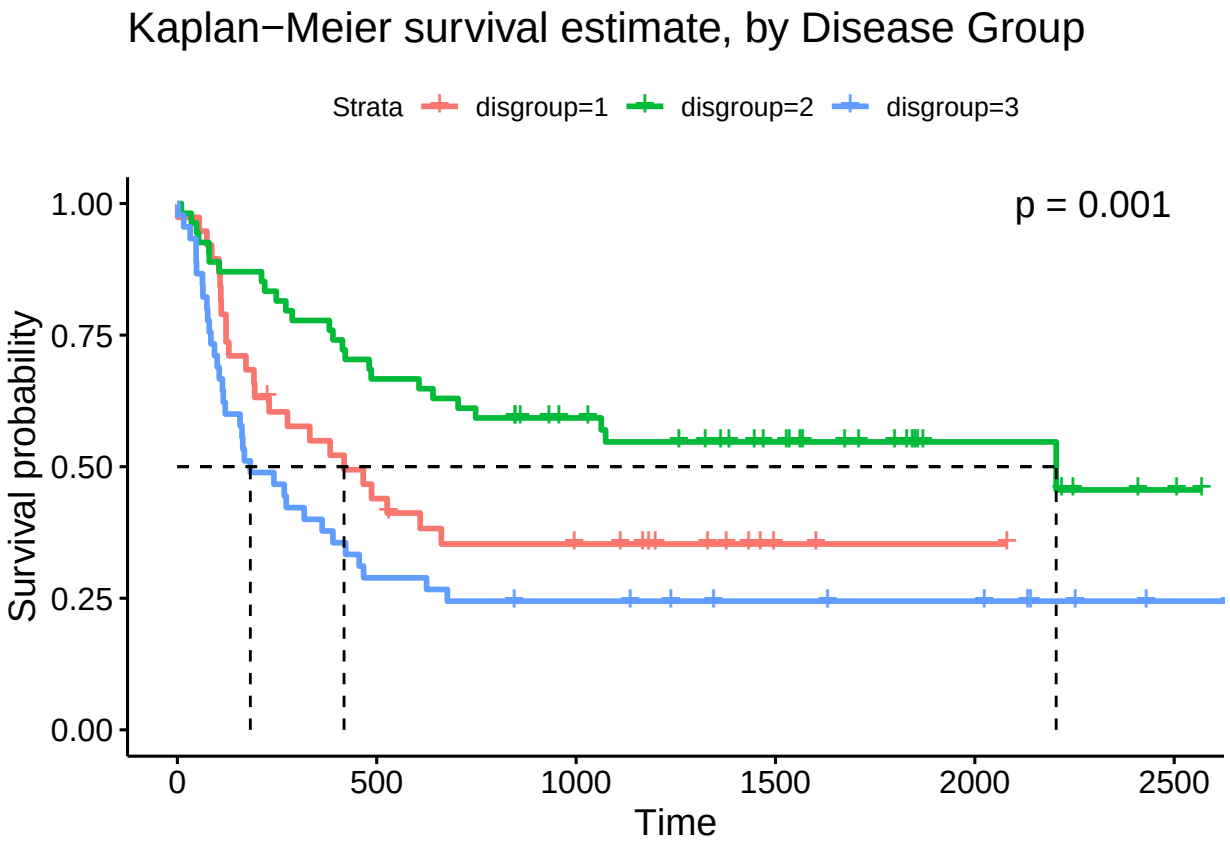



```
##   records     n.max   n.start    events     rmean se(rmean)    median   0.95LCL
##  137.0000  137.0000  137.0000   83.0000 1186.1053  100.5981  481.0000  363.0000
##   0.95UCL
##  748.0000
```

Directive 2

Table 1: Call: s_bmt ~ disgroup Chisq = 13.803722 on 2 degrees of freedom, p = 0.001006

	N	Observed	Expected	(O-E)^2/E	(O-E)^2/V
disgroup=1	38	24	21.85	0.2112	0.2893
disgroup=2	54	25	39.97	5.604	11.01
disgroup=3	45	34	21.18	7.756	10.53



	Disease Group 1	Disease Group 2	Disease Group 3
mean_age	24.421	29.407	30.444
sd_age	7.295	8.764	11.220
count_males	26.000	30.000	24.000
prop_males	0.684	0.556	0.533
count_females	12.000	24.000	21.000
prop_females	0.316	0.444	0.467
count_cmv	15.000	26.000	27.000
prop_cmv	0.395	0.481	0.600
count_mtx	17.000	12.000	11.000
prop_mtx	0.447	0.222	0.244
count_hospital	64.000	118.000	81.000
mean_donor_age	26.789	28.074	29.933
sd_donor_age	8.933	9.245	12.057
count_donor_males	26.000	34.000	28.000
prop_donor_males	0.684	0.630	0.622
count_donor_cmv	17.000	22.000	19.000
prop_donor_cmv	0.447	0.407	0.422

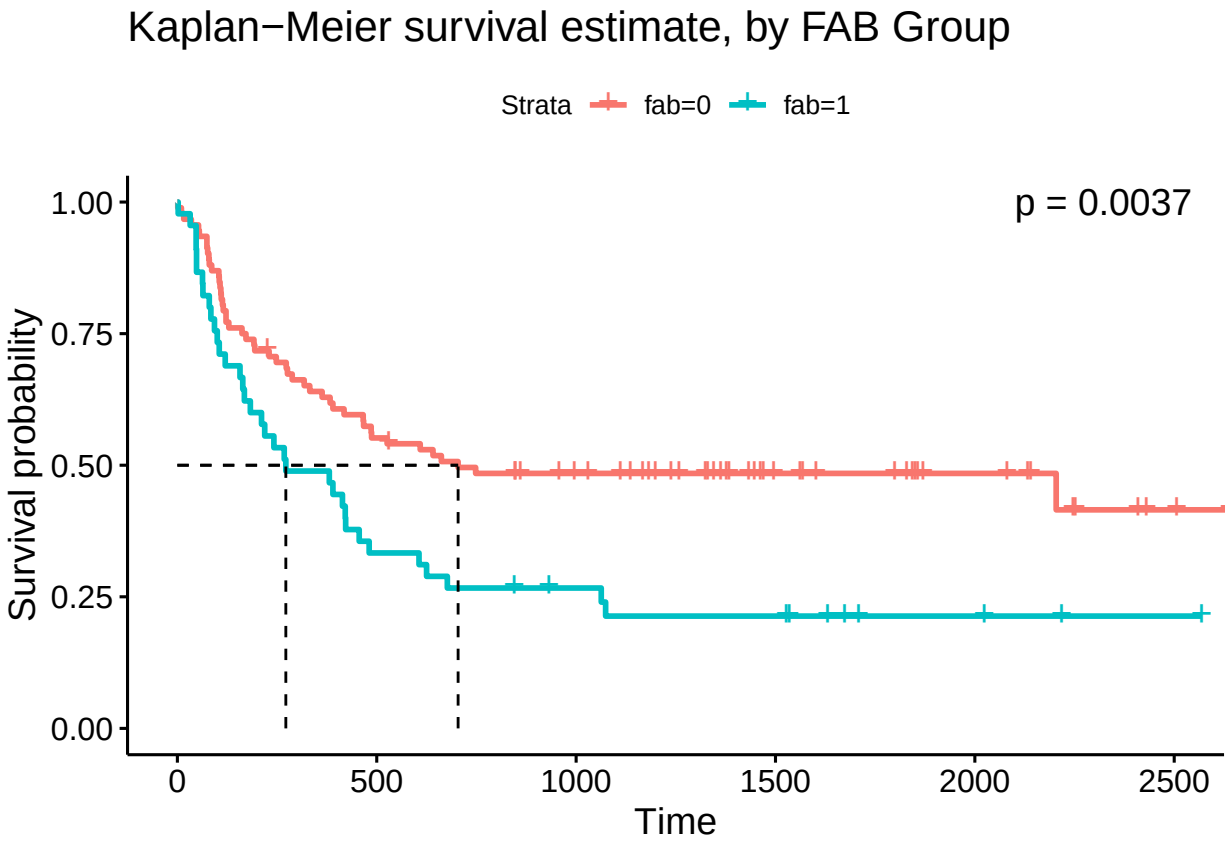
```
##               chiSq df pChisq
## 1           13.8037  2      5
## n           16.2407  2      1
## sqrtN       15.6529  2      4
## S1          15.7260  2      3
## S2          15.7781  2      2
## FH_p=1_q=1   9.9331  2      6
## $tft
##                        Q         Var        Z pNorm
## 1              -10.6695     42.7801 -1.63127     5
## n            -1294.0000 439987.8847 -1.95081     1
## sqrtN         -118.1769   4202.2583 -1.82302     4
## S1              -9.2667     23.2023 -1.92379     3
## S2              -9.1996     22.7588 -1.92839     2
## FH_p=1_q=1      -1.0948      1.4957 -0.89516     6
##
## $scores
## [1] 1 2 3
```

```
## [1] 0.05107965
```

FAB Subgrouping

Table 2: Call: s_bmt ~ fab Chisq = 8.435337 on 1 degrees of freedom, p = 0.003680

	N	Observed	Expected	(O-E)^2/E	(O-E)^2/V
fab=0	92	48	59.83	2.337	8.435
fab=1	45	35	23.17	6.034	8.435



	FAB Classification 1	FAB Classification 2
mean_age	28.598	27.889
sd_age	9.478	9.810
count_males	56.000	24.000
prop_males	0.609	0.533
count_females	36.000	21.000
prop_females	0.391	0.467
count_cmv	44.000	24.000
prop_cmv	0.478	0.533
count_mtx	32.000	8.000
prop_mtx	0.348	0.178
count_hospital	178.000	85.000
mean_donor_age	29.000	26.956
sd_donor_age	9.669	11.133
count_donor_males	58.000	30.000
prop_donor_males	0.630	0.667
count_donor_cmv	44.000	14.000
prop_donor_cmv	0.478	0.311

```
##                    Q        Var      Z pNorm
## 1          1.1825e+01 1.6590e+01 2.9033     1
## n          1.0830e+03 1.6628e+05 2.6559     6
## sqrtN      1.1217e+02 1.6047e+03 2.8001     2
## S1         7.9035e+00 8.7832e+00 2.6668     4
## S2         7.8227e+00 8.6118e+00 2.6657     5
## FH_p=1_q=1 2.1652e+00 6.0024e-01 2.7948     3
```

```
##                maxAbsZ        Var      Q pSupBr
## 1          1.2047e+01 1.6590e+01 2.9578      1
## n          1.0850e+03 1.6628e+05 2.6608      6
## sqrtN      1.1283e+02 1.6047e+03 2.8167      3
## S1         7.9834e+00 8.7832e+00 2.6938      4
## S2         7.8946e+00 8.6118e+00 2.6902      5
## FH_p=1_q=1 2.2184e+00 6.0024e-01 2.8633      2
```

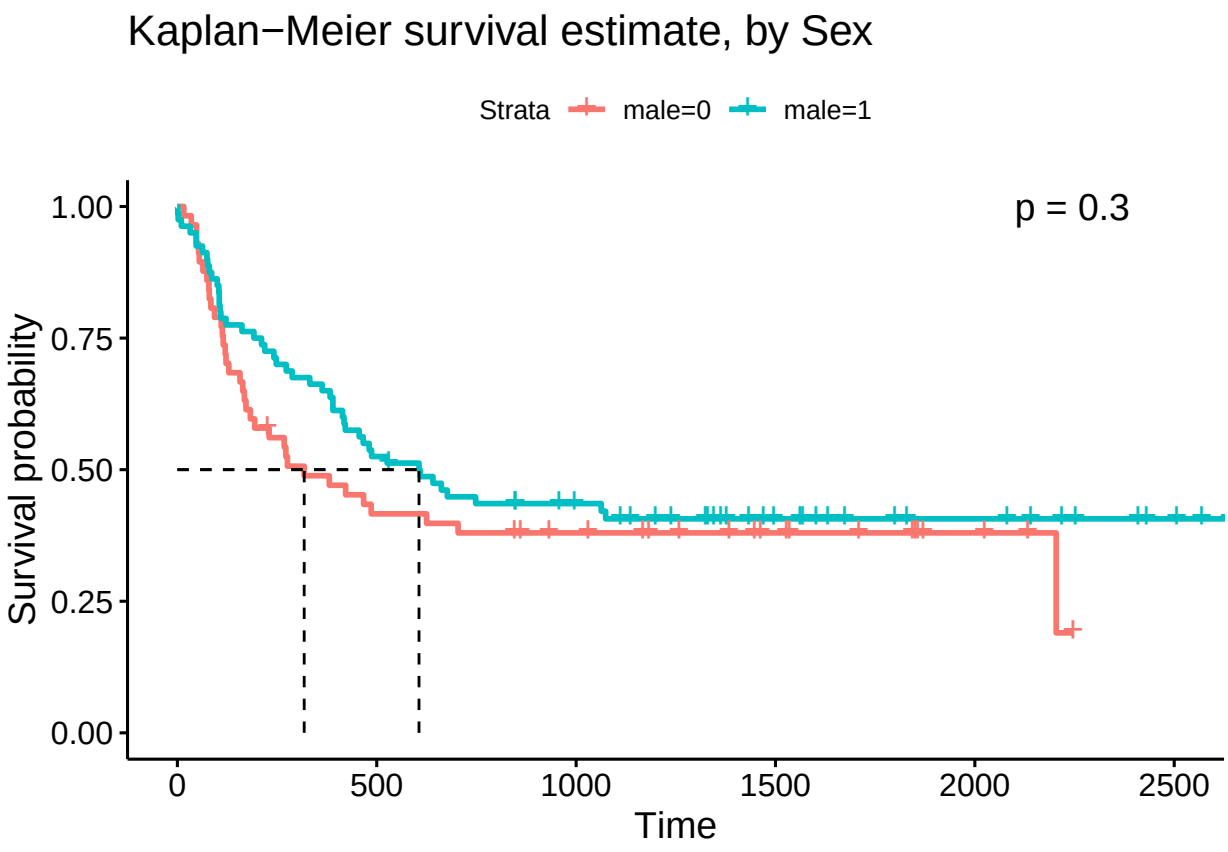
```
## [1] 0.007909707
```

Directive 3

Sex Subgrouping

Table 3: Call: s_bmt ~ male Chisq = 1.078766 on 1 degrees of freedom, p = 0.298974

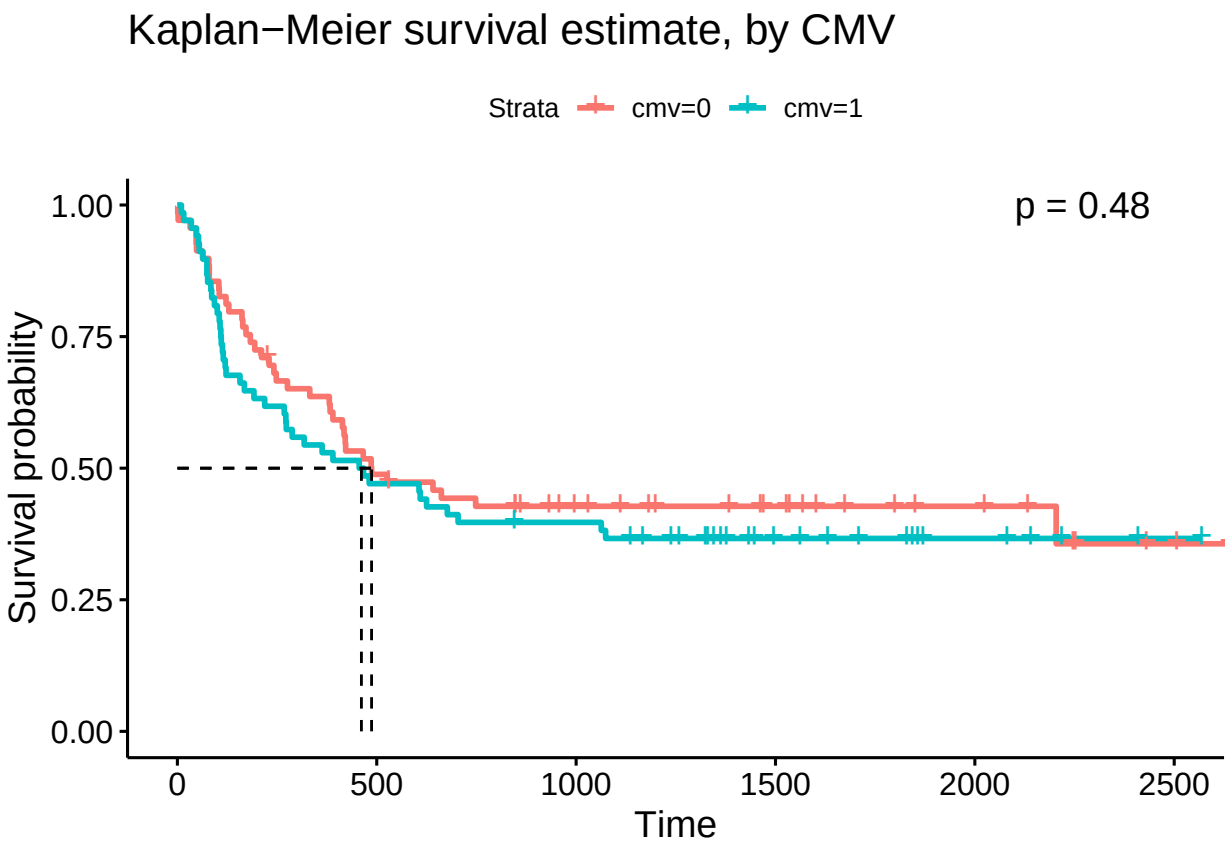
	N	Observed	Expected	(O-E)^2/E	(O-E)^2/V
male=0	57	36	31.42	0.6662	1.079
male=1	80	47	51.58	0.4059	1.079



CMV Subgrouping

Table 4: Call: s_bmt ~ cmv Chisq = 0.497423 on 1 degrees of freedom, p = 0.480635

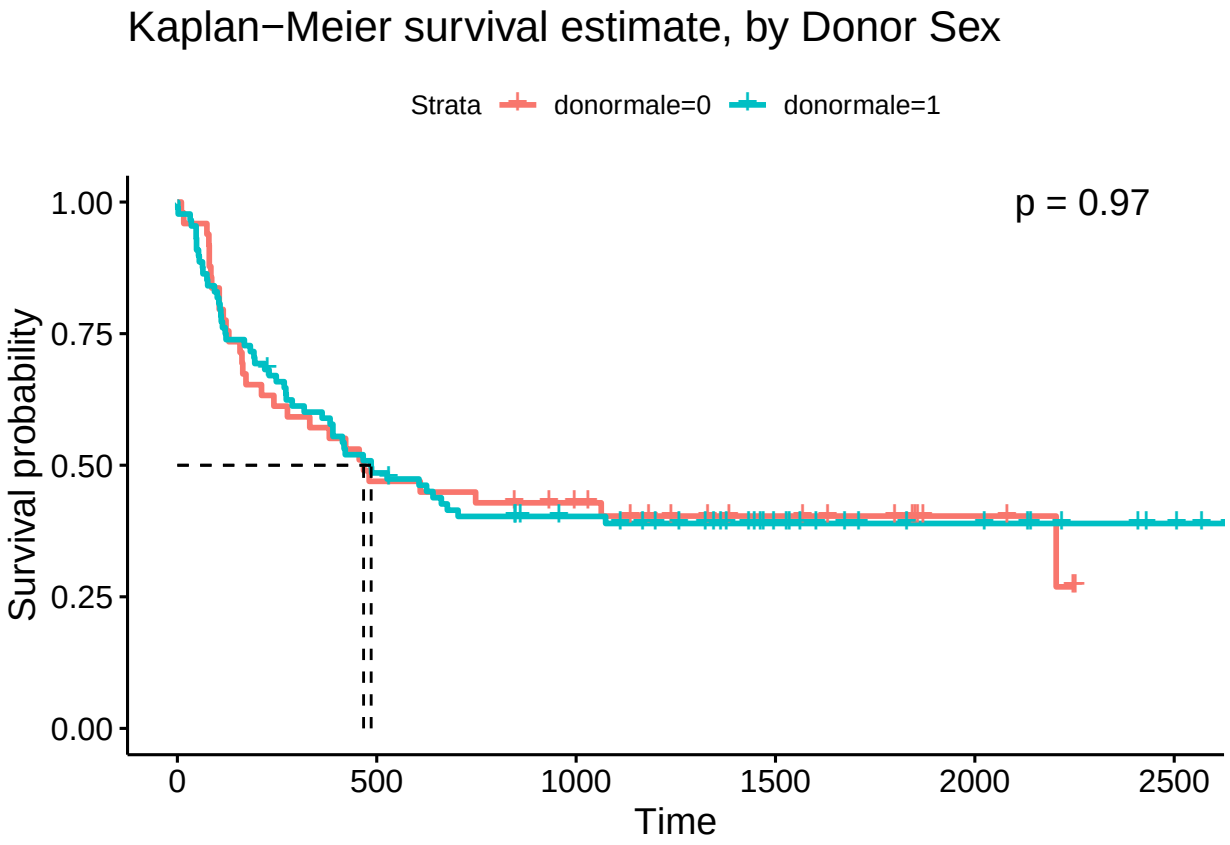
	N	Observed	Expected	(O-E)^2/E	(O-E)^2/V
cmv=0	69	40	43.2	0.2375	0.4974
cmv=1	68	43	39.8	0.2579	0.4974



Donor Sex Subgrouping

Table 5: Call: s_bmt ~ donormale Chisq = 0.001359 on 1 degrees of freedom, p = 0.970591

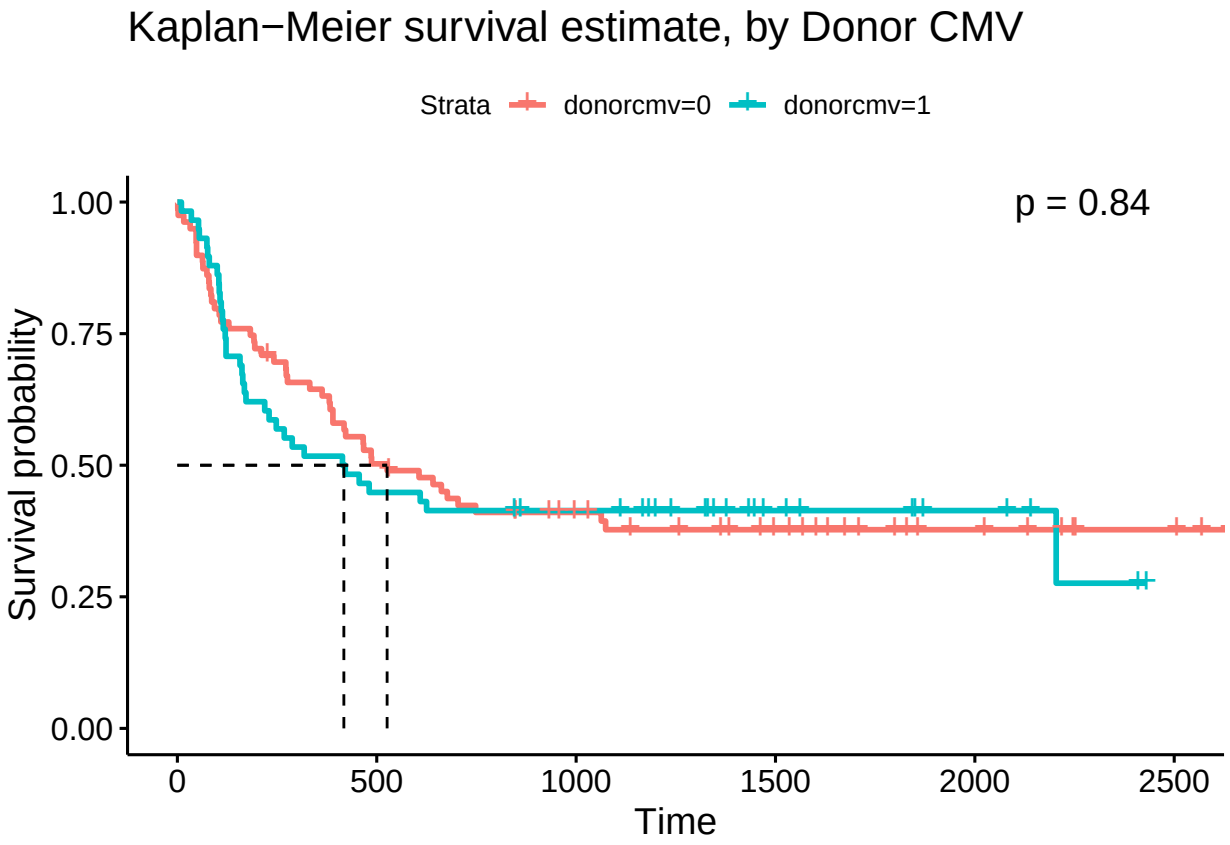
	N	Observed	Expected	(O-E)^2/E	(O-E)^2/V
donormale=0	49	30	29.84	0.0008686	0.001359
donormale=1	88	53	53.16	0.0004875	0.001359



Donor CMV Subgrouping

Table 6: Call: s_bmt ~ donorcmv Chisq = 0.043347 on 1 degrees of freedom, p = 0.835073

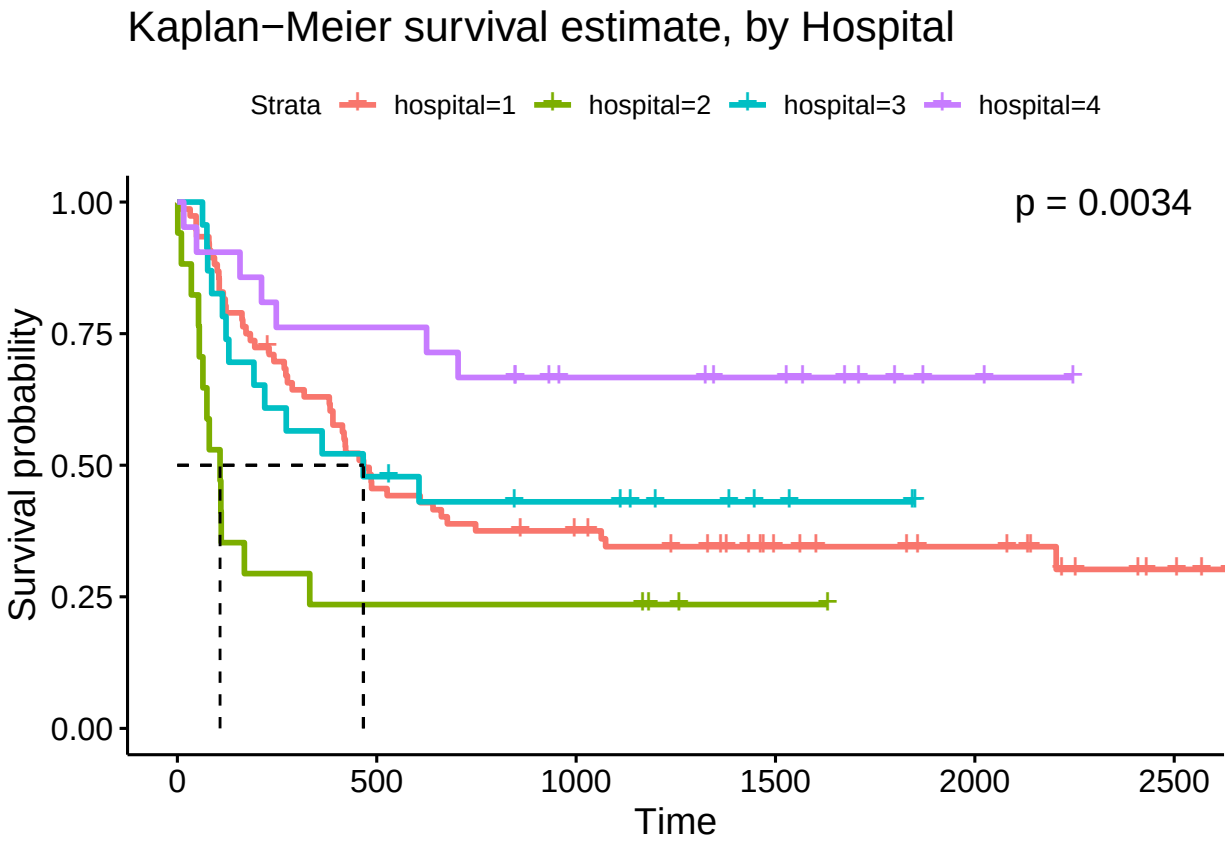
	N	Observed	Expected	(O-E)^2/E	(O-E)^2/V
donorcmv=0	79	48	48.93	0.01772	0.04335
donorcmv=1	58	35	34.07	0.02544	0.04335



Hospital Subgrouping

Table 7: Call: s_bmt ~ hospital Chisq = 13.680494 on 3 degrees of freedom, p = 0.003374

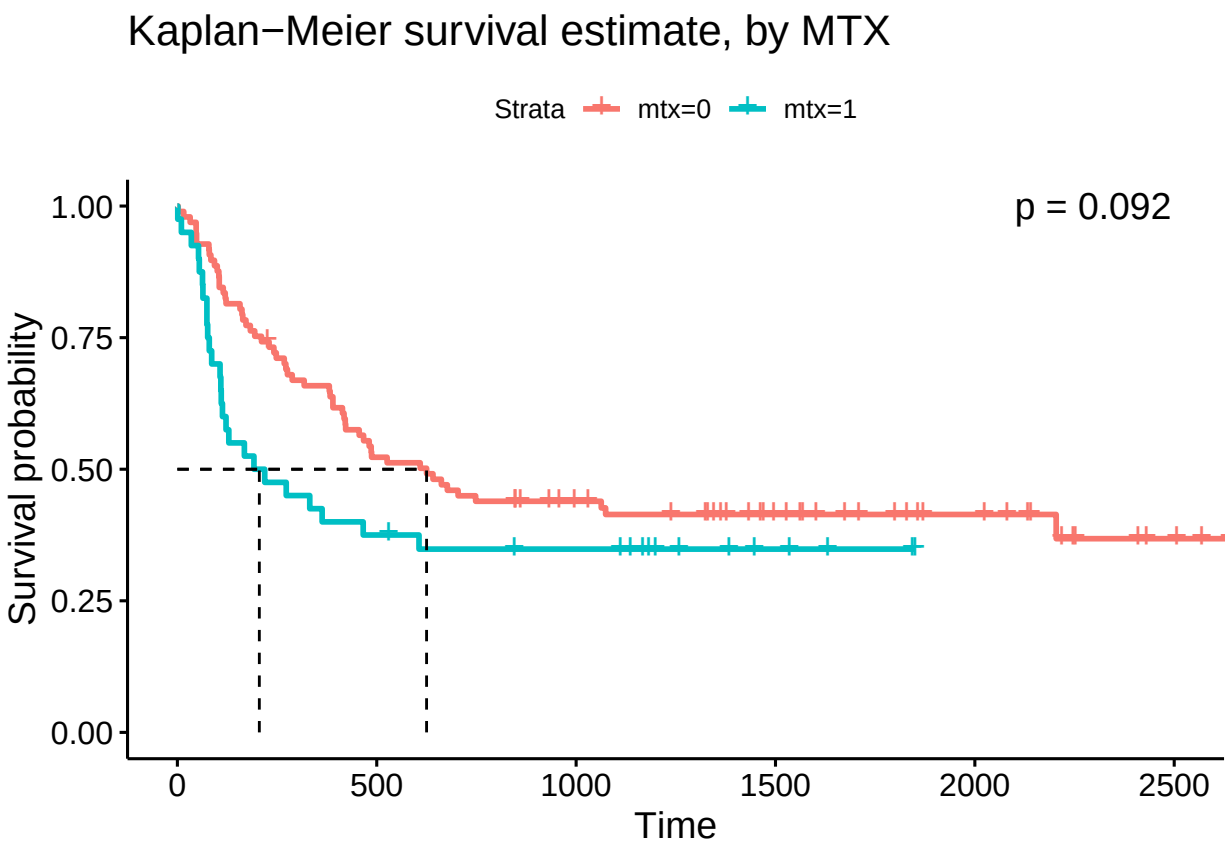
	N	Observed	Expected	(O-E)^2/E	(O-E)^2/V
hospital=1	76	50	47.71	0.1101	0.2613
hospital=2	17	13	5.905	8.524	9.258
hospital=3	23	13	13.62	0.02779	0.03339
hospital=4	21	7	15.77	4.879	6.076



MTX Subgrouping

Table 8: Call: s_bmt ~ mtx Chisq = 2.838053 on 1 degrees of freedom, p = 0.092056

	N	Observed	Expected	(O-E)^2/E	(O-E)^2/V
mtx=0	97	57	63.48	0.6614	2.838
mtx=1	40	26	19.52	2.151	2.838



Two of the hospitals appear to be significant, but since each has little data, the power might be low, the SE might be high, and the CI may cross.

#nonparametric survival function survfit.bmt <- survfit(s.bmt~1, data=bmt, conf.type="log-log")#1. estimate median disease free survival time print(survfit.bmt)#2How do patients in different disease groups or in different FAB classifications compare to each other with respect to other available baseline measurements? #table 1: columns = disease groups, rows= baseline characteristics #table 2: columns = FAB classifications, rows = baseline characteristcs#3 Are any of the measured baseline variables associated with differences in disease-free survival?#4 Is occurrence of aGVHD after transplantation associated with improved disease-free survival? summary(coxph(s.bmt ~deltaa + age + cmv + donorcmv + strata(hospital), data=bmt))#Is it associated with a decreased risk of relapse? summary(coxph(s.relapse~deltaa + age + cmv + donorcmv, data=bmt))#5 Among the patients who develop aGVHD, are any of the measured baseline factors associated with differences in disease-free survival? gvhd <- survfit(s.gvhd ~ mtx, data = bmt, conf.type = "log-log") plot(gvhd, conf.int = F, main="Kaplan-Meier GVHD survival estimate, by MTX", xlab="Time (in days)", ylab="Survival probability", col="black", lty="solid", lwd=2)gvhdcmv <- survfit(s.gvhd ~ cmv, data = bmt, conf.type = "log-log") plot(gvhdcmv, conf.int = F, main="Kaplan-Meier GVHD survival estimate, by recipient CMV status", xlab="Time (in days)", ylab="Survival probability", col="black", lty="solid", lwd=2)gvhdhospital <- survfit(s.gvhd ~ hospital, data = bmt, conf.type = "log-log") plot(gvhdhospital, conf.int = F, main="Kaplan-Meier GVHD survival estimate, by hospital", xlab="Time (in days)", ylab="Survival probability", col="black", lty="solid", lwd=2)gvhddonorcmv <- survfit(s.gvhd ~ donorcmv, data = bmt, conf.type = "log-log") plot(gvhddonorcmv, conf.int = F, main="Kaplan-Meier GVHD survival estimate, by donor CMV status", xlab="Time (in days)", ylab="Survival probability", col="black", lty="solid", lwd=2)#6 Is prophylactic use of methotrexate associated with an increased or decreased risk of developing aGVHD? **incude confounders s.gvhd <- with(bmt, Surv(ta, deltaa==1)) summary(coxph(s.gvhd~mtx + donorcmv + strata(hospital), data=bmt))#Provide an estimate of the survival function of time from transplant until onset of aGVHD separatefor patients either administered methotrexate or not. In doing so, consider the importance of accounting for relevant

confounding factors. s.gvhdm <- with(bmt, Surv(ta, deltaa==1)) survfit.gvhdmtx <- survfit(s.gvhdm~mtx, data=bmt, conf.type="log-log") plot(survfit.gvhdmtx) summary(survfit.gvhdmtx, times=c(7, 14, 21, 28, 35, 42, 49, 56))#7 Is recovery of normal platelet levels associated with improved disease-free survival? - yes summary(coxph(s.bmt ~deltap + age + donorcmv + strata(hospital), data=bmt))#Is it associated with a decreased risk of relapse? - no s.relapse <- with(bmt, Surv(agediagnosis, ageevent, deltar==1)) summary(coxph(s.relapse~deltap + age + donorcmv + strata(hospital), data=bmt)