BMT Data Analysis

Dhiraj Khanna

27 September 2016

This analysis is with the new data.

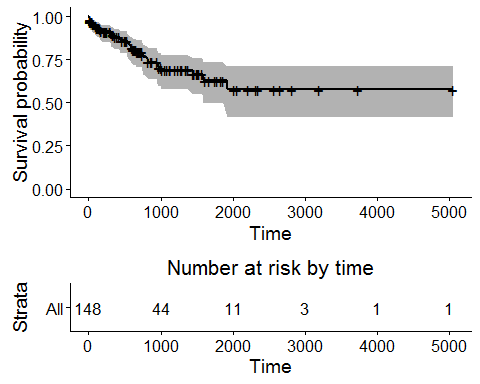
## Analysis on Death Date

### Overall

library(dplyr)  
library(ggplot2)  
library(survival)  
library(lubridate)  
library(tidyr)  
library(broom)  
library(survminer)  
  
#Read in the data  
BMT <- read.csv("BMT\_Latest.csv",na.strings=c("", "NA"))  
  
#Select variables of interest  
BMT1 <- select(BMT, TypeOfMM, Transplant.Date, C.Status,Death.date, Last.Follow.up,M.protein..,RenalFailure,Plasmacytoma,RelapseDate,ISS,TotalNoTherapy,RemissionStatusPreTx,RemissionStatusPostTx)  
BMT1$Transplant.Date <- dmy(BMT1$Transplant.Date)  
BMT1$Last.Follow.up <- dmy(BMT1$Last.Follow.up)  
BMT1$Death.date <- dmy(BMT1$Death.date)  
BMT1$RelapseDate <- dmy(BMT1$RelapseDate)  
  
#Create censoring indicator variable for death  
cenin\_dead <- rep(1, nrow(BMT1))  
for (i in 1:nrow(BMT1)) ifelse(BMT1$C.Status[i]!='Dead',cenin\_dead[i] <- 0,cenin\_dead[i] <- 1)  
  
#Replace 'NA' death dates with LFU dates  
index <- which(is.na(BMT1$Death.date))  
BMT1$Death.date[index] <- BMT1$Last.Follow.up[index]  
  
#Find time difference  
BMT1 <- mutate(BMT1, TD=as.numeric(Death.date-Transplant.Date))  
surv\_data <- cbind(BMT1,cenin\_dead)

Now let's plot the KM survival distribution for overall death.

fit\_overall <- survfit(Surv(surv\_data$TD,surv\_data$cenin\_dead)~1,conf.type = "log-log")  
ggsurvplot(fit\_overall,risk.table = T, risk.table.height = 0.4)



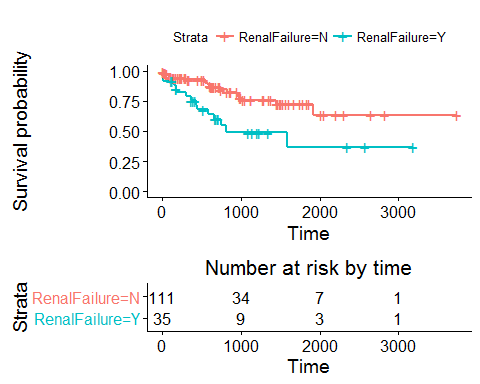
We can have a look at the model.

summary(fit\_overall)

## Call: survfit(formula = Surv(surv\_data$TD, surv\_data$cenin\_dead) ~   
## 1, conf.type = "log-log")  
##   
## 58 observations deleted due to missingness   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 9 148 1 0.993 0.00673 0.953 0.999  
## 11 147 1 0.986 0.00949 0.947 0.997  
## 15 146 1 0.980 0.01158 0.938 0.993  
## 16 144 1 0.973 0.01335 0.929 0.990  
## 21 142 1 0.966 0.01491 0.920 0.986  
## 48 130 1 0.959 0.01655 0.910 0.981  
## 89 127 1 0.951 0.01806 0.900 0.976  
## 92 126 1 0.944 0.01943 0.890 0.971  
## 112 123 1 0.936 0.02073 0.880 0.966  
## 154 114 1 0.928 0.02211 0.869 0.961  
## 174 112 1 0.919 0.02341 0.859 0.955  
## 181 108 1 0.911 0.02470 0.848 0.949  
## 309 95 1 0.901 0.02623 0.835 0.942  
## 311 94 1 0.892 0.02765 0.823 0.935  
## 365 89 1 0.882 0.02910 0.810 0.927  
## 438 85 1 0.871 0.03055 0.797 0.920  
## 439 84 1 0.861 0.03190 0.784 0.912  
## 540 78 1 0.850 0.03334 0.770 0.904  
## 545 77 1 0.839 0.03469 0.757 0.895  
## 580 76 1 0.828 0.03595 0.743 0.887  
## 585 75 1 0.817 0.03712 0.730 0.878  
## 614 73 1 0.806 0.03826 0.717 0.869  
## 663 65 1 0.793 0.03963 0.702 0.859  
## 740 56 1 0.779 0.04138 0.685 0.848  
## 750 54 1 0.765 0.04305 0.667 0.837  
## 811 53 1 0.750 0.04459 0.650 0.826  
## 827 52 1 0.736 0.04601 0.633 0.814  
## 949 47 1 0.720 0.04762 0.614 0.801  
## 969 46 1 0.704 0.04909 0.596 0.789  
## 1004 43 1 0.688 0.05061 0.577 0.775  
## 1447 29 1 0.664 0.05414 0.546 0.758  
## 1584 19 1 0.629 0.06155 0.496 0.736  
## 1917 12 1 0.577 0.07553 0.416 0.708

### Renal Failure

RF <- na.omit(select(surv\_data, RenalFailure, TD, cenin\_dead))  
fit\_RF <- survfit(Surv(TD,cenin\_dead)~RenalFailure,data=RF)  
ggsurvplot(fit\_RF,risk.table = T,risk.table.height = 0.4)

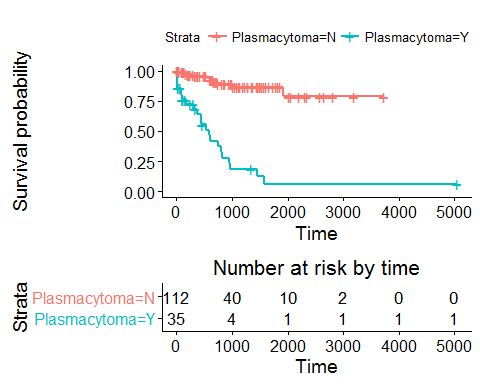


summary(fit\_RF)

## Call: survfit(formula = Surv(TD, cenin\_dead) ~ RenalFailure, data = RF)  
##   
## RenalFailure=N   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 11 111 1 0.991 0.00897 0.974 1.000  
## 21 108 1 0.982 0.01274 0.957 1.000  
## 48 96 1 0.972 0.01620 0.940 1.000  
## 89 93 1 0.961 0.01910 0.924 0.999  
## 92 92 1 0.951 0.02156 0.909 0.994  
## 112 89 1 0.940 0.02382 0.894 0.988  
## 309 68 1 0.926 0.02719 0.874 0.981  
## 540 60 1 0.911 0.03081 0.852 0.973  
## 545 59 1 0.895 0.03393 0.831 0.964  
## 585 58 1 0.880 0.03669 0.811 0.955  
## 614 56 1 0.864 0.03925 0.791 0.945  
## 740 44 1 0.845 0.04300 0.764 0.933  
## 827 42 1 0.824 0.04644 0.738 0.921  
## 949 37 1 0.802 0.05024 0.709 0.907  
## 969 36 1 0.780 0.05356 0.682 0.892  
## 1004 33 1 0.756 0.05691 0.653 0.876  
## 1447 24 1 0.725 0.06266 0.612 0.859  
## 1917 8 1 0.634 0.10093 0.464 0.866  
##   
## RenalFailure=Y   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 9 35 1 0.971 0.0282 0.918 1.000  
## 15 34 1 0.943 0.0392 0.869 1.000  
## 16 33 1 0.914 0.0473 0.826 1.000  
## 154 30 1 0.884 0.0547 0.783 0.998  
## 174 29 1 0.853 0.0607 0.742 0.981  
## 181 26 1 0.821 0.0666 0.700 0.962  
## 311 25 1 0.788 0.0716 0.659 0.941  
## 365 24 1 0.755 0.0758 0.620 0.919  
## 438 20 1 0.717 0.0808 0.575 0.894  
## 439 19 1 0.679 0.0849 0.532 0.868  
## 580 17 1 0.639 0.0889 0.487 0.840  
## 663 16 1 0.599 0.0918 0.444 0.809  
## 750 11 1 0.545 0.0983 0.383 0.776  
## 811 10 1 0.490 0.1025 0.326 0.739  
## 1584 4 1 0.368 0.1311 0.183 0.740

### Plasmacytoma

PC <- na.omit(select(surv\_data, Plasmacytoma, TD, cenin\_dead))  
fit\_PC <- survfit(Surv(TD,cenin\_dead)~Plasmacytoma,data=PC)  
ggsurvplot(fit\_PC,risk.table = T, risk.table.height = 0.4)

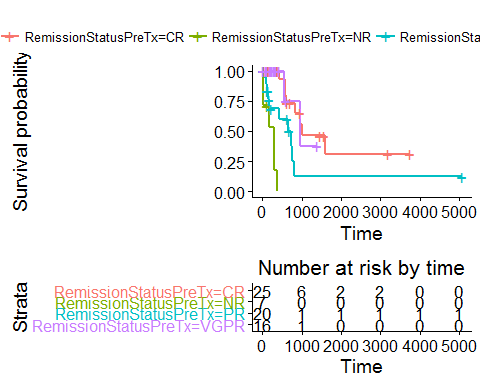


summary(fit\_PC)

## Call: survfit(formula = Surv(TD, cenin\_dead) ~ Plasmacytoma, data = PC)  
##   
## Plasmacytoma=N   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 48 101 1 0.990 0.00985 0.971 1.000  
## 154 91 1 0.979 0.01456 0.951 1.000  
## 181 87 1 0.968 0.01823 0.933 1.000  
## 309 76 1 0.955 0.02200 0.913 0.999  
## 545 66 1 0.941 0.02599 0.891 0.993  
## 585 65 1 0.926 0.02935 0.871 0.986  
## 663 56 1 0.910 0.03316 0.847 0.977  
## 740 47 1 0.890 0.03768 0.820 0.967  
## 1004 39 1 0.868 0.04308 0.787 0.956  
## 1917 11 1 0.789 0.08478 0.639 0.974  
##   
## Plasmacytoma=Y   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 9 35 1 0.9714 0.0282 0.9178 1.000  
## 11 34 1 0.9429 0.0392 0.8690 1.000  
## 15 33 1 0.9143 0.0473 0.8261 1.000  
## 16 32 1 0.8857 0.0538 0.7863 0.998  
## 21 31 1 0.8571 0.0591 0.7487 0.981  
## 89 27 1 0.8254 0.0649 0.7075 0.963  
## 92 26 1 0.7937 0.0698 0.6681 0.943  
## 112 25 1 0.7619 0.0738 0.6301 0.921  
## 174 21 1 0.7256 0.0787 0.5866 0.898  
## 311 18 1 0.6853 0.0840 0.5389 0.872  
## 365 16 1 0.6425 0.0890 0.4897 0.843  
## 438 15 1 0.5996 0.0928 0.4427 0.812  
## 439 14 1 0.5568 0.0956 0.3977 0.780  
## 540 12 1 0.5104 0.0982 0.3500 0.744  
## 580 11 1 0.4640 0.0997 0.3046 0.707  
## 614 10 1 0.4176 0.0999 0.2613 0.667  
## 750 9 1 0.3712 0.0990 0.2201 0.626  
## 811 8 1 0.3248 0.0969 0.1810 0.583  
## 827 7 1 0.2784 0.0935 0.1441 0.538  
## 949 6 1 0.2320 0.0887 0.1097 0.491  
## 969 5 1 0.1856 0.0822 0.0779 0.442  
## 1447 3 1 0.1237 0.0745 0.0380 0.403  
## 1584 2 1 0.0619 0.0575 0.0100 0.382

### Remission Status Before

RSPre <- na.omit(select(surv\_data, RemissionStatusPreTx, TD, cenin\_dead))  
fit\_RSPre <- survfit(Surv(TD,cenin\_dead)~RemissionStatusPreTx,data=RSPre)  
ggsurvplot(fit\_RSPre,risk.table = T,risk.table.height = 0.4)

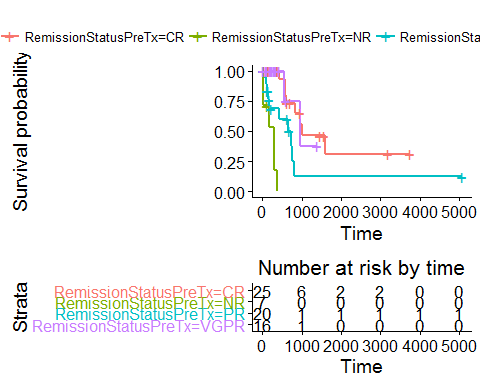


summary(fit\_RSPre)

## Call: survfit(formula = Surv(TD, cenin\_dead) ~ RemissionStatusPreTx,   
## data = RSPre)  
##   
## RemissionStatusPreTx=CR   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 438 15 1 0.933 0.0644 0.815 1.000  
## 580 14 1 0.867 0.0878 0.711 1.000  
## 585 13 1 0.800 0.1033 0.621 1.000  
## 614 12 1 0.733 0.1142 0.540 0.995  
## 827 9 1 0.652 0.1273 0.445 0.956  
## 969 7 1 0.559 0.1391 0.343 0.910  
## 1004 6 1 0.466 0.1437 0.254 0.853  
## 1584 3 1 0.310 0.1589 0.114 0.846  
##   
## RemissionStatusPreTx=NR   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 11 7 1 0.857 0.132 0.6334 1  
## 15 6 1 0.714 0.171 0.4471 1  
## 174 4 1 0.536 0.201 0.2570 1  
## 309 3 1 0.357 0.198 0.1205 1  
## 311 2 1 0.179 0.160 0.0307 1  
## 365 1 1 0.000 NaN NA NA  
##   
## RemissionStatusPreTx=PR   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 48 18 1 0.944 0.0540 0.8443 1.000  
## 89 17 1 0.889 0.0741 0.7549 1.000  
## 112 16 1 0.833 0.0878 0.6778 1.000  
## 154 12 1 0.764 0.1044 0.5843 0.999  
## 181 10 1 0.688 0.1187 0.4902 0.964  
## 439 8 1 0.602 0.1313 0.3922 0.923  
## 663 6 1 0.501 0.1427 0.2870 0.876  
## 740 4 1 0.376 0.1524 0.1699 0.832  
## 750 3 1 0.251 0.1442 0.0812 0.774  
## 811 2 1 0.125 0.1142 0.0210 0.748  
##   
## RemissionStatusPreTx=VGPR   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 545 4 1 0.750 0.217 0.4259 1  
## 949 2 1 0.375 0.286 0.0839 1

### Remission Status After

RSPost <- na.omit(select(surv\_data, RemissionStatusPostTx, TD, cenin\_dead))  
fit\_RSPost <- survfit(Surv(TD,cenin\_dead)~RemissionStatusPostTx,data=RSPost)  
ggsurvplot(fit\_RSPre,risk.table = T, risk.table.height = 0.4)

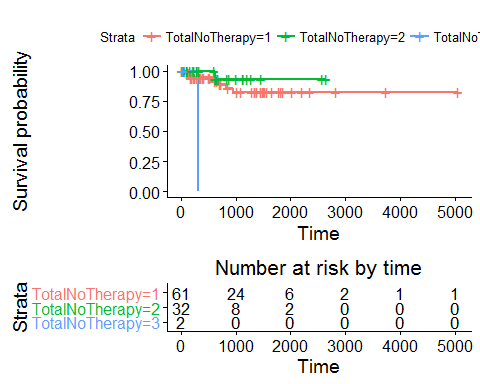


summary(fit\_RSPost)

## Call: survfit(formula = Surv(TD, cenin\_dead) ~ RemissionStatusPostTx,   
## data = RSPost)  
##   
## RemissionStatusPostTx=   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
##   
## RemissionStatusPostTx=CR   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 154 73 1 0.986 0.0136 0.960 1.000  
## 174 71 1 0.972 0.0192 0.935 1.000  
## 545 51 1 0.953 0.0267 0.902 1.000  
## 585 50 1 0.934 0.0323 0.873 1.000  
## 827 37 1 0.909 0.0401 0.834 0.991  
## 949 32 1 0.881 0.0478 0.792 0.980  
## 1917 9 1 0.783 0.1016 0.607 1.000  
##   
## RemissionStatusPostTx=Neg   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
##   
## RemissionStatusPostTx=NR   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 11 4 1 0.75 0.217 0.426 1  
## 15 3 1 0.50 0.250 0.188 1  
## 365 1 1 0.00 NaN NA NA  
##   
## RemissionStatusPostTx=PR   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 48 11 1 0.909 0.0867 0.7541 1.000  
## 112 10 1 0.818 0.1163 0.6192 1.000  
## 181 9 1 0.727 0.1343 0.5064 1.000  
## 311 8 1 0.636 0.1450 0.4071 0.995  
## 614 7 1 0.545 0.1501 0.3180 0.936  
## 663 6 1 0.455 0.1501 0.2379 0.868  
## 740 4 1 0.341 0.1495 0.1443 0.805  
## 750 3 1 0.227 0.1362 0.0702 0.736  
##   
## RemissionStatusPostTx=VGPR   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI

### Number of Therapies

thp <- na.omit(select(surv\_data, TotalNoTherapy, TD, cenin\_dead))  
fit\_thp <- survfit(Surv(TD,cenin\_dead)~TotalNoTherapy,data=thp)  
ggsurvplot(fit\_thp,risk.table = T, risk.table.height = 0.4)



summary(fit\_thp)

## Call: survfit(formula = Surv(TD, cenin\_dead) ~ TotalNoTherapy, data = thp)  
##   
## TotalNoTherapy=1   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 89 58 1 0.983 0.0171 0.950 1.000  
## 154 52 1 0.964 0.0251 0.916 1.000  
## 174 51 1 0.945 0.0309 0.886 1.000  
## 585 36 1 0.919 0.0397 0.844 1.000  
## 663 32 1 0.890 0.0477 0.801 0.989  
## 827 27 1 0.857 0.0562 0.754 0.975  
## 949 25 1 0.823 0.0635 0.707 0.957  
##   
## TotalNoTherapy=2   
## time n.risk n.event survival std.err   
## 614.0000 15.0000 1.0000 0.9333 0.0644   
## lower 95% CI upper 95% CI   
## 0.8153 1.0000   
##   
## TotalNoTherapy=3   
## time n.risk n.event survival std.err   
## 309 1 1 0 NaN   
## lower 95% CI upper 95% CI   
## NA NA

### ISS

Values 'II' and 'III' have been coerced to a single value - 'II'.

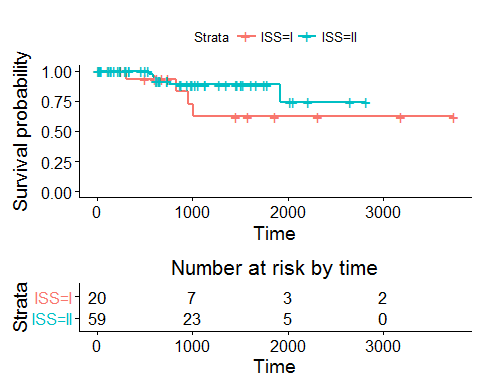
iss <- na.omit(select(surv\_data, ISS, TD, cenin\_dead))  
iss[iss=='III',] <- "II"  
iss$ISS <- factor(iss$ISS)  
iss$TD <- as.numeric(iss$TD)

## Warning: NAs introduced by coercion

iss$cenin\_dead <- as.numeric(iss$cenin\_dead)

## Warning: NAs introduced by coercion

fit\_iss <- survfit(Surv(TD,cenin\_dead)~ISS,data=iss)  
ggsurvplot(fit\_iss,risk.table = T, risk.table.height = 0.4)



summary(fit\_thp)

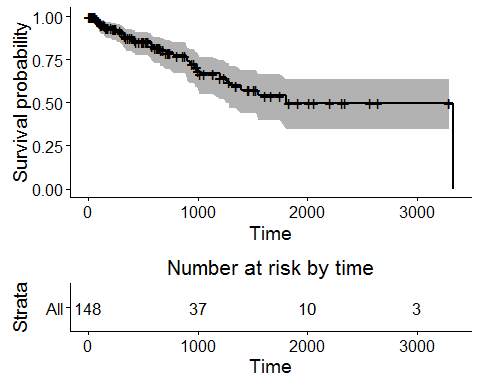
## Call: survfit(formula = Surv(TD, cenin\_dead) ~ TotalNoTherapy, data = thp)  
##   
## TotalNoTherapy=1   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 89 58 1 0.983 0.0171 0.950 1.000  
## 154 52 1 0.964 0.0251 0.916 1.000  
## 174 51 1 0.945 0.0309 0.886 1.000  
## 585 36 1 0.919 0.0397 0.844 1.000  
## 663 32 1 0.890 0.0477 0.801 0.989  
## 827 27 1 0.857 0.0562 0.754 0.975  
## 949 25 1 0.823 0.0635 0.707 0.957  
##   
## TotalNoTherapy=2   
## time n.risk n.event survival std.err   
## 614.0000 15.0000 1.0000 0.9333 0.0644   
## lower 95% CI upper 95% CI   
## 0.8153 1.0000   
##   
## TotalNoTherapy=3   
## time n.risk n.event survival std.err   
## 309 1 1 0 NaN   
## lower 95% CI upper 95% CI   
## NA NA

## Analysis on Relapse Date

#Create censoring indicator variable for relapse date  
relapse <- rep(1, nrow(surv\_data))  
for (i in 1:nrow(surv\_data)) ifelse(is.na(surv\_data$RelapseDate[i]),relapse[i] <- 0,relapse[i] <- 1)  
  
#Replace 'NA' death dates with LFU dates  
index <- which(is.na(surv\_data$RelapseDate))  
surv\_data$RelapseDate[index] <- surv\_data$Last.Follow.up[index]  
  
#time difference for transplant - relapse  
surv\_data <- mutate(surv\_data, TD\_R=as.numeric(RelapseDate-Transplant.Date))  
rlp\_data <- cbind(surv\_data,relapse)

Now that the data is ready, we can do a similar analysis. Start off with an overall analysis.

fit\_overall <- survfit(Surv(rlp\_data$TD\_R,rlp\_data$relapse)~1,conf.type = "log-log")  
ggsurvplot(fit\_overall,risk.table = T, risk.table.height = 0.4)



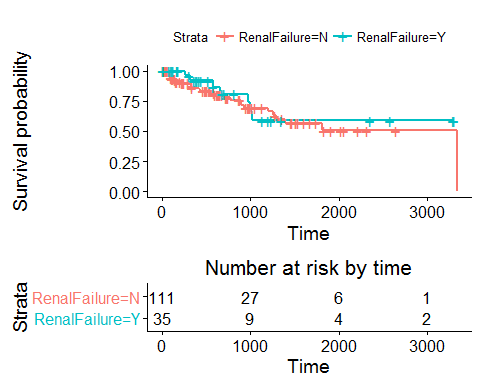
We can have a look at the model.

summary(fit\_overall)

## Call: survfit(formula = Surv(rlp\_data$TD\_R, rlp\_data$relapse) ~ 1,   
## conf.type = "log-log")  
##   
## 58 observations deleted due to missingness   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 67 129 1 0.992 0.00772 0.946 0.999  
## 76 128 1 0.984 0.01088 0.939 0.996  
## 83 126 1 0.977 0.01330 0.929 0.992  
## 94 124 1 0.969 0.01535 0.919 0.988  
## 99 122 1 0.961 0.01716 0.909 0.984  
## 112 119 1 0.953 0.01882 0.898 0.979  
## 145 113 1 0.944 0.02045 0.887 0.973  
## 157 111 1 0.936 0.02197 0.876 0.967  
## 216 104 1 0.927 0.02353 0.864 0.961  
## 263 95 1 0.917 0.02522 0.851 0.955  
## 306 93 1 0.907 0.02681 0.838 0.948  
## 315 91 1 0.897 0.02831 0.826 0.941  
## 329 90 1 0.887 0.02970 0.813 0.933  
## 339 87 1 0.877 0.03106 0.800 0.926  
## 416 82 1 0.866 0.03247 0.787 0.918  
## 426 80 1 0.856 0.03382 0.774 0.909  
## 570 70 1 0.843 0.03548 0.758 0.900  
## 574 69 1 0.831 0.03701 0.743 0.891  
## 596 66 1 0.819 0.03853 0.728 0.881  
## 656 59 1 0.805 0.04030 0.711 0.871  
## 705 55 1 0.790 0.04214 0.693 0.860  
## 771 50 1 0.774 0.04416 0.673 0.848  
## 891 46 1 0.757 0.04629 0.652 0.835  
## 897 45 1 0.741 0.04823 0.632 0.822  
## 931 44 1 0.724 0.04998 0.612 0.808  
## 975 41 1 0.706 0.05179 0.591 0.794  
## 997 38 1 0.688 0.05365 0.569 0.779  
## 1015 36 1 0.668 0.05546 0.547 0.764  
## 1200 31 1 0.647 0.05771 0.522 0.747  
## 1257 28 1 0.624 0.06009 0.495 0.729  
## 1289 26 1 0.600 0.06239 0.467 0.710  
## 1391 23 1 0.574 0.06490 0.437 0.689  
## 1549 17 1 0.540 0.06930 0.396 0.664  
## 1806 13 1 0.498 0.07539 0.344 0.635  
## 3335 1 1 0.000 NaN NA NA

### Renal Failure

RF <- na.omit(select(rlp\_data, RenalFailure, TD\_R, relapse))  
fit\_RF <- survfit(Surv(TD\_R,relapse)~RenalFailure,data=RF)  
ggsurvplot(fit\_RF,risk.table = T,risk.table.height = 0.4)

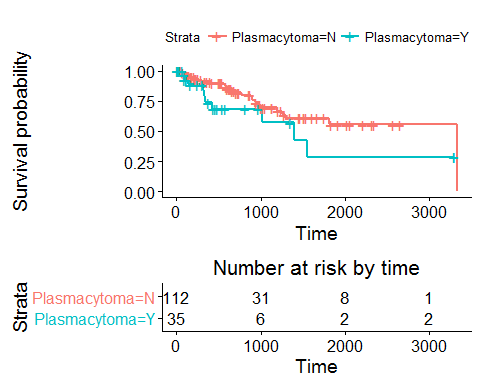


summary(fit\_RF)

## Call: survfit(formula = Surv(TD\_R, relapse) ~ RenalFailure, data = RF)  
##   
## RenalFailure=N   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 67 94 1 0.989 0.0106 0.969 1.000  
## 76 93 1 0.979 0.0149 0.950 1.000  
## 83 92 1 0.968 0.0181 0.933 1.000  
## 94 90 1 0.957 0.0209 0.917 0.999  
## 99 89 1 0.947 0.0232 0.902 0.993  
## 112 86 1 0.936 0.0255 0.887 0.987  
## 145 82 1 0.924 0.0276 0.872 0.980  
## 157 80 1 0.913 0.0296 0.856 0.972  
## 216 76 1 0.901 0.0315 0.841 0.965  
## 306 66 1 0.887 0.0339 0.823 0.956  
## 329 65 1 0.873 0.0360 0.806 0.947  
## 339 63 1 0.859 0.0380 0.788 0.937  
## 416 61 1 0.845 0.0399 0.771 0.927  
## 426 60 1 0.831 0.0416 0.754 0.917  
## 570 52 1 0.815 0.0438 0.734 0.906  
## 596 50 1 0.799 0.0459 0.714 0.894  
## 705 41 1 0.779 0.0487 0.690 0.881  
## 771 37 1 0.758 0.0517 0.663 0.867  
## 891 34 1 0.736 0.0548 0.636 0.852  
## 897 33 1 0.714 0.0575 0.610 0.836  
## 931 32 1 0.691 0.0599 0.584 0.819  
## 1200 23 1 0.661 0.0644 0.547 0.800  
## 1257 22 1 0.631 0.0681 0.511 0.780  
## 1289 20 1 0.600 0.0717 0.475 0.758  
## 1391 18 1 0.566 0.0750 0.437 0.734  
## 1806 9 1 0.504 0.0893 0.356 0.713  
## 3335 1 1 0.000 NaN NA NA  
##   
## RenalFailure=Y   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 263 26 1 0.962 0.0377 0.890 1.000  
## 315 24 1 0.921 0.0533 0.823 1.000  
## 574 17 1 0.867 0.0727 0.736 1.000  
## 656 15 1 0.809 0.0879 0.654 1.000  
## 975 11 1 0.736 0.1063 0.554 0.977  
## 997 10 1 0.662 0.1185 0.466 0.940  
## 1015 9 1 0.589 0.1261 0.387 0.896

### Plasmacytoma

PC <- na.omit(select(rlp\_data, Plasmacytoma, TD\_R, relapse))  
fit\_PC <- survfit(Surv(TD\_R,relapse)~Plasmacytoma,data=PC)  
ggsurvplot(fit\_PC,risk.table = T, risk.table.height = 0.4)

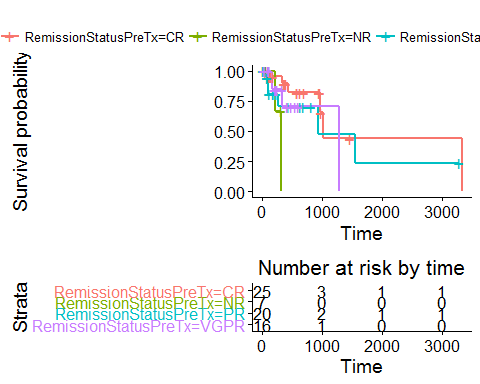


summary(fit\_PC)

## Call: survfit(formula = Surv(TD\_R, relapse) ~ Plasmacytoma, data = PC)  
##   
## Plasmacytoma=N   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 67 100 1 0.990 0.00995 0.971 1.000  
## 76 99 1 0.980 0.01400 0.953 1.000  
## 99 97 1 0.970 0.01712 0.937 1.000  
## 112 94 1 0.960 0.01980 0.922 0.999  
## 157 89 1 0.949 0.02232 0.906 0.994  
## 216 83 1 0.937 0.02481 0.890 0.987  
## 263 75 1 0.925 0.02745 0.873 0.980  
## 306 73 1 0.912 0.02985 0.856 0.973  
## 416 68 1 0.899 0.03229 0.838 0.964  
## 570 61 1 0.884 0.03496 0.818 0.955  
## 574 60 1 0.869 0.03735 0.799 0.946  
## 596 58 1 0.854 0.03960 0.780 0.936  
## 656 51 1 0.838 0.04222 0.759 0.925  
## 705 47 1 0.820 0.04492 0.736 0.913  
## 771 42 1 0.800 0.04791 0.712 0.900  
## 891 39 1 0.780 0.05088 0.686 0.886  
## 897 38 1 0.759 0.05352 0.661 0.872  
## 931 37 1 0.739 0.05587 0.637 0.857  
## 975 35 1 0.718 0.05812 0.612 0.841  
## 997 32 1 0.695 0.06048 0.586 0.824  
## 1200 26 1 0.668 0.06379 0.554 0.806  
## 1257 23 1 0.639 0.06731 0.520 0.786  
## 1289 21 1 0.609 0.07066 0.485 0.764  
## 1806 11 1 0.554 0.08314 0.412 0.743  
## 3335 1 1 0.000 NaN NA NA  
##   
## Plasmacytoma=Y   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 83 28 1 0.964 0.0351 0.8979 1.000  
## 94 26 1 0.927 0.0496 0.8349 1.000  
## 145 22 1 0.885 0.0627 0.7702 1.000  
## 315 18 1 0.836 0.0761 0.6992 0.999  
## 329 17 1 0.787 0.0861 0.6349 0.975  
## 339 16 1 0.738 0.0937 0.5750 0.946  
## 426 14 1 0.685 0.1007 0.5133 0.914  
## 1015 6 1 0.571 0.1338 0.3605 0.904  
## 1391 4 1 0.428 0.1592 0.2065 0.887  
## 1549 3 1 0.285 0.1576 0.0967 0.842

### Remission Status Before

RSPre <- na.omit(select(rlp\_data, RemissionStatusPreTx, TD\_R, relapse))  
fit\_RSPre <- survfit(Surv(TD\_R,relapse)~RemissionStatusPreTx,data=RSPre)  
ggsurvplot(fit\_RSPre,risk.table = T, risk.table.height = 0.4)

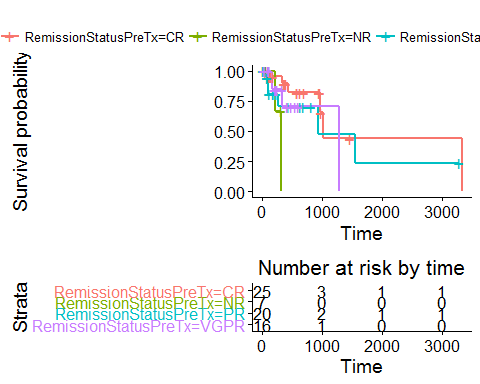


summary(fit\_RSPre)

## Call: survfit(formula = Surv(TD\_R, relapse) ~ RemissionStatusPreTx,   
## data = RSPre)  
##   
## RemissionStatusPreTx=CR   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 67 22 1 0.955 0.0444 0.871 1  
## 329 16 1 0.895 0.0712 0.766 1  
## 426 12 1 0.820 0.0967 0.651 1  
## 975 5 1 0.656 0.1659 0.400 1  
## 1015 3 1 0.438 0.2101 0.171 1  
## 3335 1 1 0.000 NaN NA NA  
##   
## RemissionStatusPreTx=NR   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 216 3 1 0.667 0.272 0.3 1  
## 315 1 1 0.000 NaN NA NA  
##   
## RemissionStatusPreTx=PR   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 76 17 1 0.941 0.0571 0.8357 1  
## 94 14 1 0.874 0.0837 0.7244 1  
## 99 13 1 0.807 0.1007 0.6316 1  
## 263 8 1 0.706 0.1291 0.4933 1  
## 931 3 1 0.471 0.2105 0.1958 1  
## 1549 2 1 0.235 0.1969 0.0456 1  
##   
## RemissionStatusPreTx=VGPR   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 145 13 1 0.923 0.0739 0.789 1  
## 157 12 1 0.846 0.1001 0.671 1  
## 339 6 1 0.705 0.1534 0.460 1  
## 1289 1 1 0.000 NaN NA NA

### Remission Status After

RSPost <- na.omit(select(rlp\_data, RemissionStatusPostTx, TD\_R, relapse))  
fit\_RSPost <- survfit(Surv(TD\_R,relapse)~RemissionStatusPostTx,data=RSPost)  
ggsurvplot(fit\_RSPre,risk.table = T, risk.table.height = 0.4)

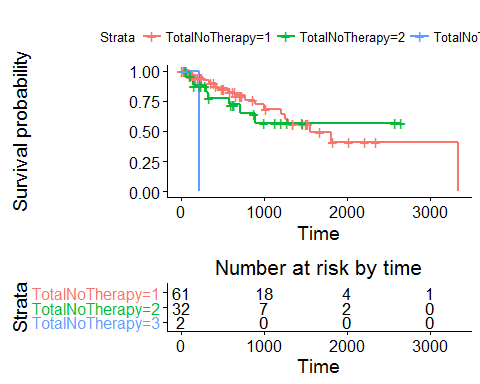


summary(fit\_RSPost)

## Call: survfit(formula = Surv(TD\_R, relapse) ~ RemissionStatusPostTx,   
## data = RSPost)  
##   
## RemissionStatusPostTx=   
## time n.risk n.event survival std.err   
## 416 1 1 0 NaN   
## lower 95% CI upper 95% CI   
## NA NA   
##   
## RemissionStatusPostTx=CR   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 67 81 1 0.988 0.0123 0.964 1.000  
## 76 80 1 0.975 0.0172 0.942 1.000  
## 99 77 1 0.963 0.0212 0.922 1.000  
## 112 76 1 0.950 0.0244 0.903 0.999  
## 145 71 1 0.937 0.0275 0.884 0.992  
## 306 57 1 0.920 0.0315 0.860 0.984  
## 426 51 1 0.902 0.0357 0.835 0.975  
## 570 46 1 0.883 0.0399 0.808 0.964  
## 574 45 1 0.863 0.0436 0.782 0.953  
## 596 43 1 0.843 0.0470 0.756 0.940  
## 656 38 1 0.821 0.0507 0.727 0.926  
## 705 36 1 0.798 0.0542 0.698 0.911  
## 771 32 1 0.773 0.0579 0.667 0.895  
## 891 29 1 0.746 0.0618 0.635 0.878  
## 897 28 1 0.720 0.0651 0.603 0.859  
## 975 26 1 0.692 0.0682 0.570 0.839  
## 997 25 1 0.664 0.0709 0.539 0.819  
## 1200 19 1 0.629 0.0753 0.498 0.796  
## 1549 10 1 0.566 0.0903 0.414 0.774  
## 1806 8 1 0.496 0.1031 0.330 0.745  
## 3335 1 1 0.000 NaN NA NA  
##   
## RemissionStatusPostTx=Neg   
## time n.risk n.event survival std.err   
## 1257 1 1 0 NaN   
## lower 95% CI upper 95% CI   
## NA NA   
##   
## RemissionStatusPostTx=NR   
## time n.risk n.event survival std.err   
## 315 1 1 0 NaN   
## lower 95% CI upper 95% CI   
## NA NA   
##   
## RemissionStatusPostTx=PR   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 94 10 1 0.900 0.0949 0.732 1  
## 263 8 1 0.787 0.1340 0.564 1  
## 329 6 1 0.656 0.1638 0.402 1  
## 931 4 1 0.492 0.1878 0.233 1  
##   
## RemissionStatusPostTx=VGPR   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 157 3 1 0.667 0.272 0.2995 1  
## 339 2 1 0.333 0.272 0.0673 1

### Number of Therapies

thp <- na.omit(select(rlp\_data, TotalNoTherapy, TD\_R, relapse))  
fit\_thp <- survfit(Surv(TD\_R,relapse)~TotalNoTherapy,data=thp)  
ggsurvplot(fit\_thp,risk.table = T, risk.table.height = 0.4)



summary(fit\_thp)

## Call: survfit(formula = Surv(TD\_R, relapse) ~ TotalNoTherapy, data = thp)  
##   
## TotalNoTherapy=1   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 67 58 1 0.983 0.0171 0.950 1.000  
## 99 54 1 0.965 0.0246 0.917 1.000  
## 157 49 1 0.945 0.0310 0.886 1.000  
## 263 42 1 0.922 0.0376 0.852 0.999  
## 339 41 1 0.900 0.0429 0.820 0.988  
## 416 38 1 0.876 0.0478 0.787 0.975  
## 426 36 1 0.852 0.0523 0.755 0.961  
## 570 31 1 0.824 0.0574 0.719 0.945  
## 656 27 1 0.794 0.0629 0.680 0.927  
## 771 22 1 0.758 0.0696 0.633 0.907  
## 897 20 1 0.720 0.0757 0.586 0.885  
## 997 19 1 0.682 0.0807 0.541 0.860  
## 1200 17 1 0.642 0.0853 0.495 0.833  
## 1257 16 1 0.602 0.0889 0.450 0.804  
## 1289 15 1 0.562 0.0916 0.408 0.773  
## 1549 8 1 0.491 0.1036 0.325 0.743  
## 1806 6 1 0.410 0.1142 0.237 0.707  
## 3335 1 1 0.000 NaN NA NA  
##   
## TotalNoTherapy=2   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 76 26 1 0.962 0.0377 0.890 1.000  
## 112 23 1 0.920 0.0545 0.819 1.000  
## 145 22 1 0.878 0.0662 0.757 1.000  
## 306 17 1 0.826 0.0799 0.684 0.999  
## 329 16 1 0.775 0.0901 0.617 0.973  
## 574 14 1 0.719 0.0992 0.549 0.943  
## 705 10 1 0.647 0.1124 0.461 0.910  
## 891 8 1 0.566 0.1241 0.369 0.870  
##   
## TotalNoTherapy=3   
## time n.risk n.event survival std.err   
## 216 1 1 0 NaN   
## lower 95% CI upper 95% CI   
## NA NA

### ISS

Values 'II' and 'III' have been coerced to a single value - 'II'.

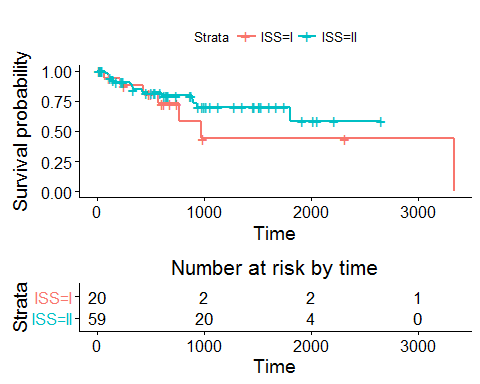
iss <- na.omit(select(rlp\_data, ISS, TD\_R, relapse))  
iss[iss=='III',] <- "II"  
iss$ISS <- factor(iss$ISS)  
iss$TD\_R <- as.numeric(iss$TD\_R)

## Warning: NAs introduced by coercion

iss$relapse <- as.numeric(iss$relapse)

## Warning: NAs introduced by coercion

fit\_iss <- survfit(Surv(TD\_R,relapse)~ISS,data=iss)  
ggsurvplot(fit\_iss,risk.table = T, risk.table.height = 0.4)



summary(fit\_thp)

## Call: survfit(formula = Surv(TD\_R, relapse) ~ TotalNoTherapy, data = thp)  
##   
## TotalNoTherapy=1   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 67 58 1 0.983 0.0171 0.950 1.000  
## 99 54 1 0.965 0.0246 0.917 1.000  
## 157 49 1 0.945 0.0310 0.886 1.000  
## 263 42 1 0.922 0.0376 0.852 0.999  
## 339 41 1 0.900 0.0429 0.820 0.988  
## 416 38 1 0.876 0.0478 0.787 0.975  
## 426 36 1 0.852 0.0523 0.755 0.961  
## 570 31 1 0.824 0.0574 0.719 0.945  
## 656 27 1 0.794 0.0629 0.680 0.927  
## 771 22 1 0.758 0.0696 0.633 0.907  
## 897 20 1 0.720 0.0757 0.586 0.885  
## 997 19 1 0.682 0.0807 0.541 0.860  
## 1200 17 1 0.642 0.0853 0.495 0.833  
## 1257 16 1 0.602 0.0889 0.450 0.804  
## 1289 15 1 0.562 0.0916 0.408 0.773  
## 1549 8 1 0.491 0.1036 0.325 0.743  
## 1806 6 1 0.410 0.1142 0.237 0.707  
## 3335 1 1 0.000 NaN NA NA  
##   
## TotalNoTherapy=2   
## time n.risk n.event survival std.err lower 95% CI upper 95% CI  
## 76 26 1 0.962 0.0377 0.890 1.000  
## 112 23 1 0.920 0.0545 0.819 1.000  
## 145 22 1 0.878 0.0662 0.757 1.000  
## 306 17 1 0.826 0.0799 0.684 0.999  
## 329 16 1 0.775 0.0901 0.617 0.973  
## 574 14 1 0.719 0.0992 0.549 0.943  
## 705 10 1 0.647 0.1124 0.461 0.910  
## 891 8 1 0.566 0.1241 0.369 0.870  
##   
## TotalNoTherapy=3   
## time n.risk n.event survival std.err   
## 216 1 1 0 NaN   
## lower 95% CI upper 95% CI   
## NA NA