Surrogate for FRET LCA Biomarker

P Barber

12 April 2019

In the tables that shows the patient characteristics for the FRET LCA and full LCA classes some covariates have a significant difference between the classes. KM curves are plotted here to see if these have any predictive value (replicating the LCA class splits).

Load data.

```
load(file = "COIN_Final.Rdata")
```

Survival Curves split by FRET high vs low, mlivonly and SUMLES, in FRET data set
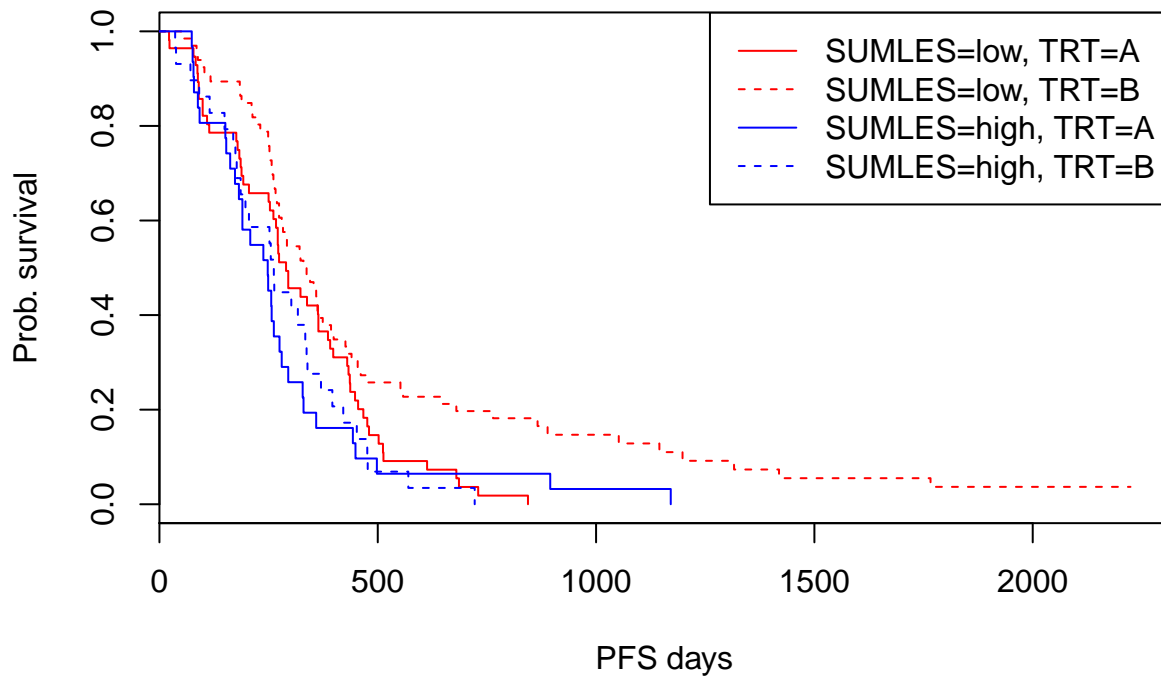
Exploring whether PIK3CA and FRET split the patients. FRET and SUMLES split by median.

The patient characteristics table and FRET_Class_Cluster_Means.Rmd/pdf tells us that Class 1 (those who respond to TRT) may be identified.

Try to identify a target Treatment group.

Test individual variables
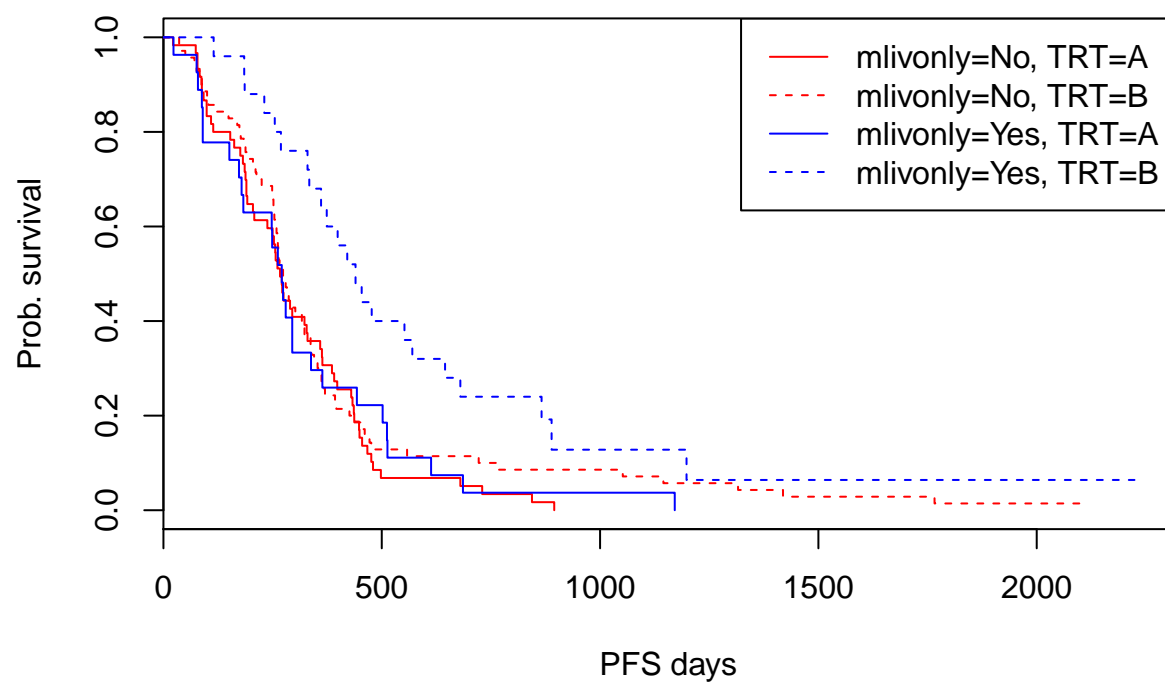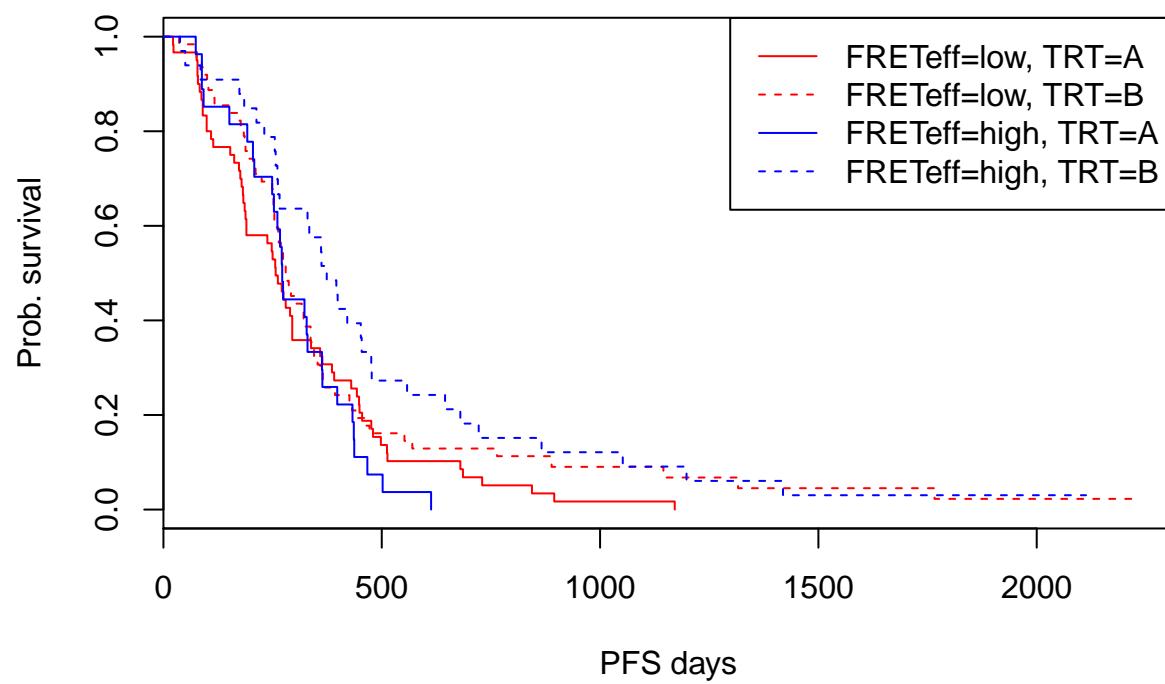
Test FRET only

```
## Call: survfit(formula = SurvObj.pfs ~ FRETeff + TRT, data = data)
##
##     9 observations deleted due to missingness
##                        n events median 0.95LCL 0.95UCL
## FRETeff=low, TRT=A   60      59    257     190     338
## FRETeff=low, TRT=B   62      59    282     259     337
## FRETeff=high, TRT=A  27      27    272     253     364
## FRETeff=high, TRT=B  33      32    374     266     477

## Call:
## survdiff(formula = SurvObj.pfs[FRETeff == "low"] ~ TRT[FRETeff ==
##     "low"], data = data)
##
## n=122, 9 observations deleted due to missingness.
##
##                            N Observed Expected (O-E)^2/E (O-E)^2/V
## TRT[FRETeff == "low"]=A 60       59     52.6     0.790      1.47
## TRT[FRETeff == "low"]=B 62       59     65.4     0.635      1.47
##
##  Chisq= 1.5  on 1 degrees of freedom, p= 0.2

## Call:
## survdiff(formula = SurvObj.pfs[FRETeff == "high"] ~ TRT[FRETeff ==
##     "high"], data = data)
```

3

```
## 
## n=60, 9 observations deleted due to missingness.
## 
##                              N Observed Expected (O-E)^2/E (O-E)^2/V
## TRT[FRETeff == "high"]=A 27        27     18.1      4.32      6.95
## TRT[FRETeff == "high"]=B 33        32     40.9      1.92      6.95
## 
##  Chisq= 7  on 1 degrees of freedom, p= 0.008
```

Test liver only

```
## Call: survfit(formula = SurvObj.pfs ~ mlivonly + TRT, data = data)
## 
##    9 observations deleted due to missingness
##                     n events median 0.95LCL 0.95UCL
## mlivonly=No, TRT=A  60     59    267     238     359
## mlivonly=No, TRT=B  70     69    274     257     337
## mlivonly=Yes, TRT=A 27     27    271     183     364
## mlivonly=Yes, TRT=B 25     22    440     361     680

## Call:
## survdiff(formula = SurvObj.pfs[mlivonly == "Yes"] ~ TRT[mlivonly ==
##     "Yes"], data = data)
## 
## n=52, 9 observations deleted due to missingness.
## 
##                                N Observed Expected (O-E)^2/E (O-E)^2/V
## TRT[mlivonly == "Yes"]=A 27          27     17.8      4.79      7.87
## TRT[mlivonly == "Yes"]=B 25          22     31.2      2.73      7.87
## 
##  Chisq= 7.9  on 1 degrees of freedom, p= 0.005
```
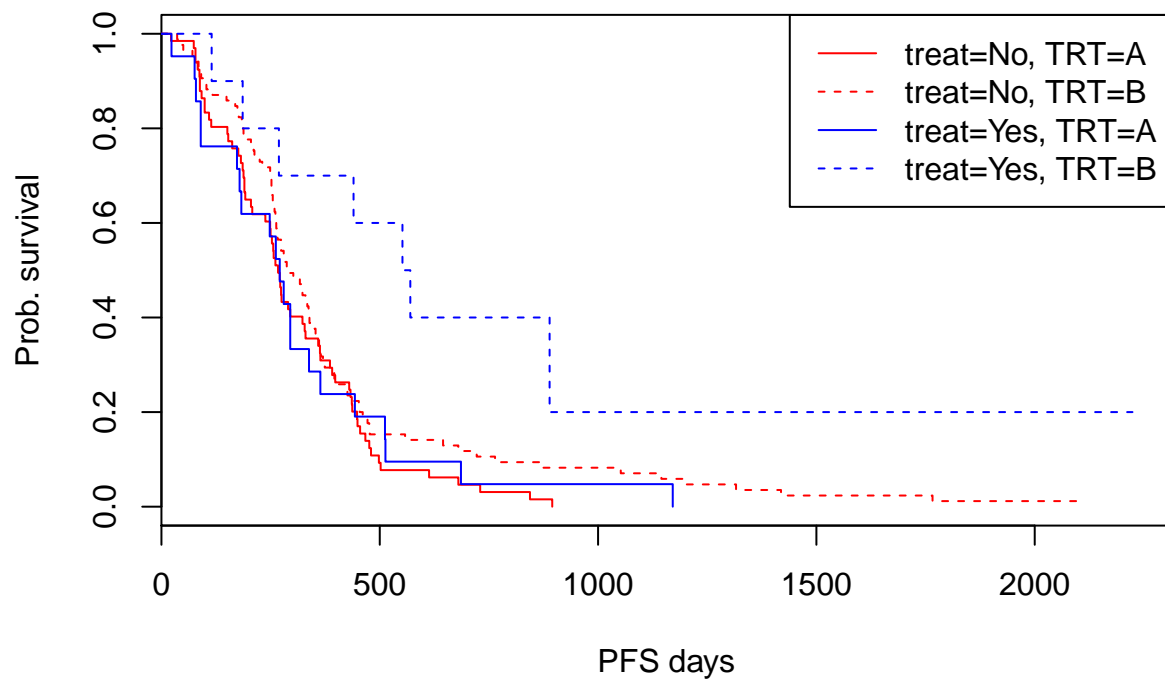
Test SUMLES only

```
## Call: survfit(formula = SurvObj.pfs ~ SUMLES + TRT, data = data)
## 
##    9 observations deleted due to missingness
##                    n events median 0.95LCL 0.95UCL
## SUMLES=low, TRT=A  56     55    290     261     386
## SUMLES=low, TRT=B  66     62    334     274     399
## SUMLES=high, TRT=A 31     31    248     182     295
## SUMLES=high, TRT=B 29     29    262     197     339

## Call:
## survdiff(formula = SurvObj.pfs[SUMLES == "low"] ~ TRT[SUMLES ==
##     "low"], data = data)
## 
## n=122, 9 observations deleted due to missingness.
## 
##                               N Observed Expected (O-E)^2/E (O-E)^2/V
## TRT[SUMLES == "low"]=A 56          55     43.1      3.29       5.6
## TRT[SUMLES == "low"]=B 66          62     73.9      1.92       5.6
## 
```

```
##  Chisq= 5.6  on 1 degrees of freedom, p= 0.02

## Call:
## survdiff(formula = SurvObj.pfs[SUMLES == "high"] ~ TRT[SUMLES ==
##     "high"], data = data)
##
## n=60, 9 observations deleted due to missingness.
##
##                            N Observed Expected (O-E)^2/E (O-E)^2/V
## TRT[SUMLES == "high"]=A 31       31     29.6    0.0618     0.129
## TRT[SUMLES == "high"]=B 29       29     30.4    0.0603     0.129
##
##  Chisq= 0.1  on 1 degrees of freedom, p= 0.7
```

Test FRET low + liver only

```
data$treat  <- ifelse(data$FRETeff=="low" & data$mlivonly=="Yes", "Yes", "No")
```
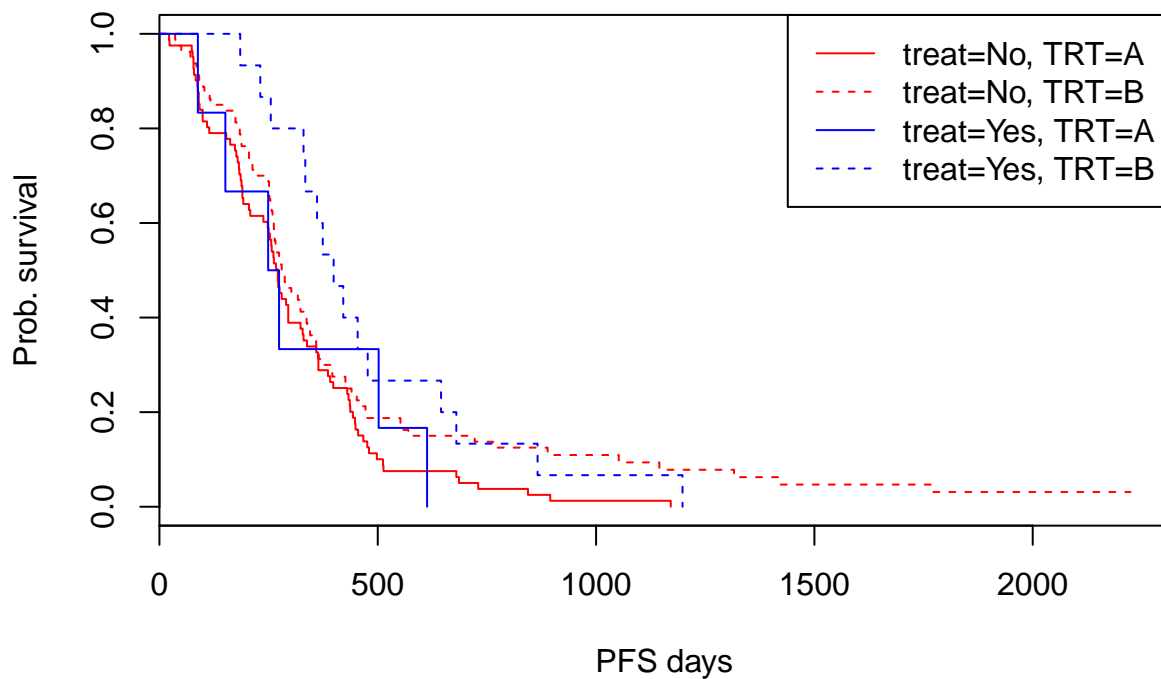


```
## Call: survfit(formula = SurvObj.pfs ~ treat + TRT, data = data)
##
##     9 observations deleted due to missingness
##                 n events median 0.95LCL 0.95UCL
## treat=No, TRT=A  66     65    267     238     330
## treat=No, TRT=B  85     84    292     263     345
## treat=Yes, TRT=A 21     21    271     179     443
## treat=Yes, TRT=B 10      7    561     269      NA
```

```
## Call:
## survdiff(formula = SurvObj.pfs[treat == "Yes"] ~ TRT[treat ==
##      "Yes"], data = data)
##
## n=31, 9 observations deleted due to missingness.
##
##                         N Observed Expected (O-E)^2/E (O-E)^2/V
## TRT[treat == "Yes"]=A 21       21     14.8      2.64      6.01
## TRT[treat == "Yes"]=B 10        7     13.2      2.94      6.01
##
##  Chisq= 6  on 1 degrees of freedom, p= 0.01
```

Test FRET high + liver only

```
data$treat    <- ifelse(data$FRETeff=="high" & data$mlivonly=="Yes", "Yes", "No")
data$treat_os <- ifelse(data$FRETeff=="high" & data$mlivonly=="Yes", "Yes", "No")
```
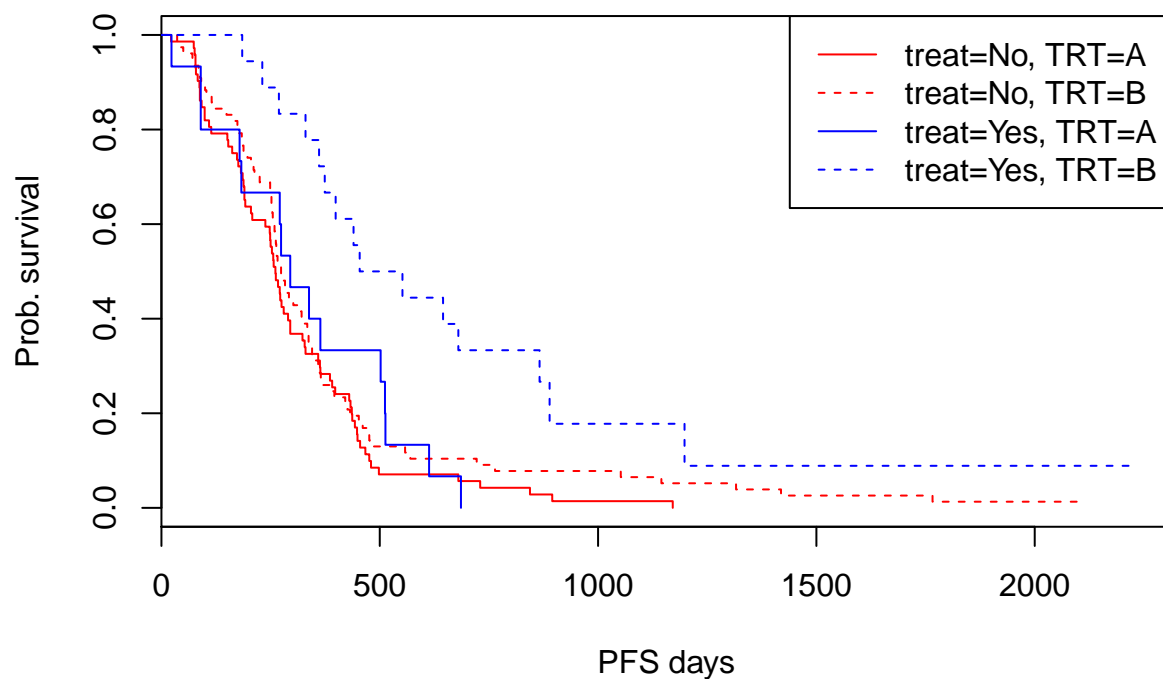


```
## Call: survfit(formula = SurvObj.pfs ~ treat + TRT, data = data)
##
##     9 observations deleted due to missingness
##                  n events median 0.95LCL 0.95UCL
## treat=No, TRT=A  81     80    271     248     323
## treat=No, TRT=B  80     76    282     262     339
## treat=Yes, TRT=A  6      6    262     151      NA
## treat=Yes, TRT=B 15     15    399     334     680
```

6

```
## Call:
## survdiff(formula = SurvObj.pfs[treat == "Yes"] ~ TRT[treat ==
##     "Yes"], data = data)
##
## n=21, 9 observations deleted due to missingness.
##
##                          N Observed Expected (O-E)^2/E (O-E)^2/V
## TRT[treat == "Yes"]=A  6        6     3.73     1.385       1.8
## TRT[treat == "Yes"]=B 15       15    17.27     0.299       1.8
##
##  Chisq= 1.8  on 1 degrees of freedom, p= 0.2
```

Test SUMLES + liver only

```
data$treat  <- ifelse(data$SUMLES=="low" & data$mlivonly=="Yes", "Yes", "No")
```
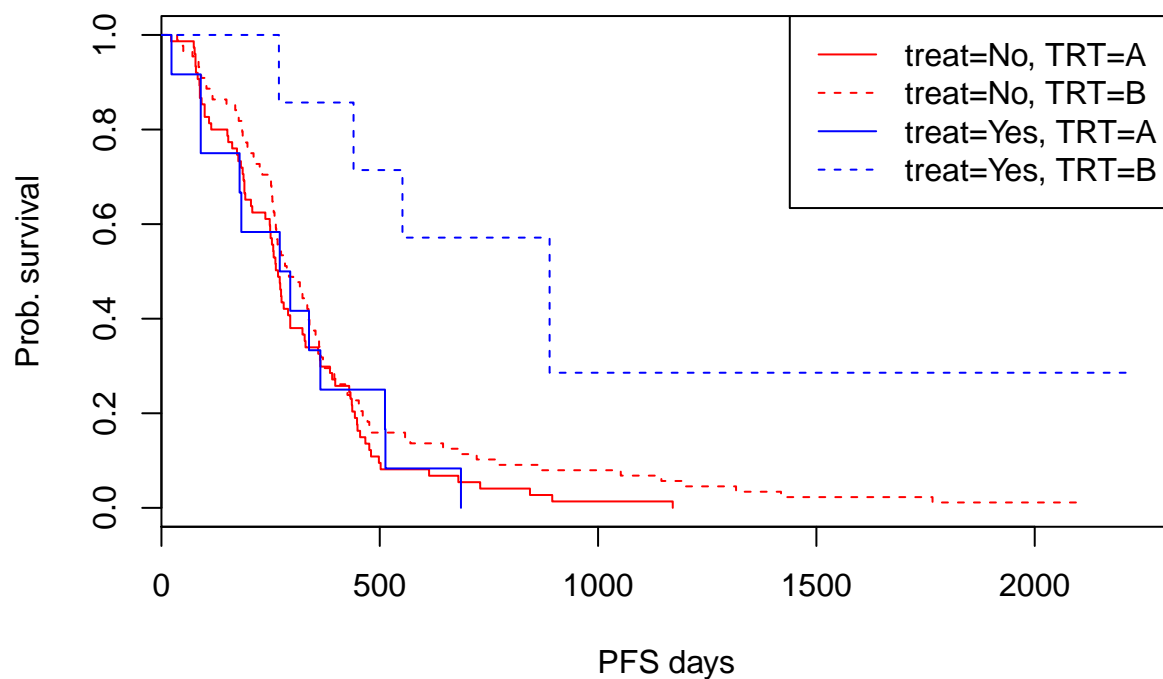


```
## Call: survfit(formula = SurvObj.pfs ~ treat + TRT, data = data)
##
##     9 observations deleted due to missingness
##                   n events median 0.95LCL 0.95UCL
## treat=No, TRT=A  72       71    261     238     295
## treat=No, TRT=B  77       76    274     257     337
## treat=Yes, TRT=A 15       15    295     183     513
## treat=Yes, TRT=B 18       15    503     374      NA
##
## Call:
```

```
## survdiff(formula = SurvObj.pfs[treat == "Yes"] ~ TRT[treat ==
##     "Yes"], data = data)
##
## n=33, 9 observations deleted due to missingness.
##
##                            N Observed Expected (O-E)^2/E (O-E)^2/V
## TRT[treat == "Yes"]=A 15          15      8.5      4.96      7.56
## TRT[treat == "Yes"]=B 18          15     21.5      1.96      7.56
##
##  Chisq= 7.6  on 1 degrees of freedom, p= 0.006
```

Test FRET + liver only + SUMLES

```r
data$treat  <- ifelse(data$FRETeff=="low" &
                        data$mlivonly=="Yes" & data$SUMLES=="low", "Yes", "No")
```


PFS days

```
## Call: survfit(formula = SurvObj.pfs ~ treat + TRT, data = data)
##
##     9 observations deleted due to missingness
##                   n events median 0.95LCL 0.95UCL
## treat=No, TRT=A  75     74    267     248     323
## treat=No, TRT=B  88     87    290     262     345
## treat=Yes, TRT=A 12     12    283     179      NA
## treat=Yes, TRT=B  7      4    889     440      NA
##
## Call:
```

```
## survdiff(formula = SurvObj.pfs[treat == "Yes"] ~ TRT[treat ==
##     "Yes"], data = data)
##
## n=19, 9 observations deleted due to missingness.
##
##                        N Observed Expected (O-E)^2/E (O-E)^2/V
## TRT[treat == "Yes"]=A 12       12      6.7      4.19      8.32
## TRT[treat == "Yes"]=B  7        4      9.3      3.02      8.32
##
##  Chisq= 8.3  on 1 degrees of freedom, p= 0.004
```

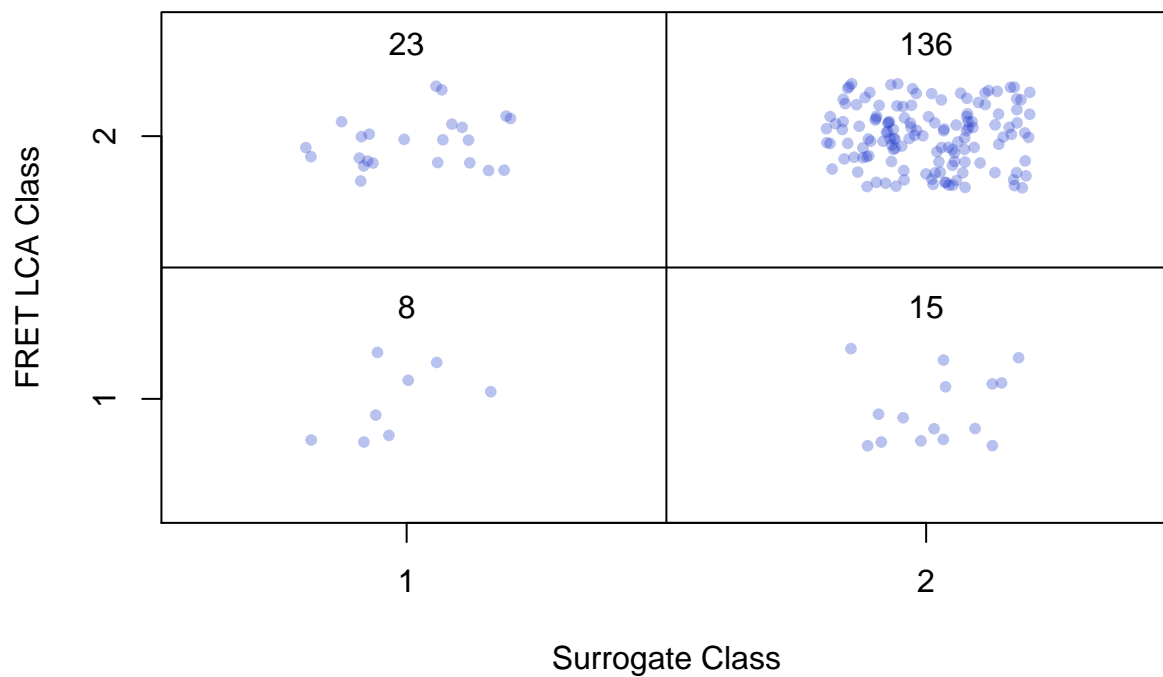## Overlap with LCA CLasses

Create classes from treat="Yes" -> Class 1.

```r
data$treat  <- ifelse(data$FRETeff=="low" & data$mlivonly=="Yes", "Yes", "No")

data$surrogate_class <- ifelse(data$treat=="Yes", 1, 2)

#data$surrogate_class[is.na(data$surrogate_class)] <- 2

as.numeric.factor <- function(x) {as.numeric(levels(x))[x]}
data$Class.FRET.PFS <- as.numeric.factor(data$Class.FRET.PFS)
```

Cohen's Kappa

```
library(irr)
```

```
## Warning: package 'irr' was built under R version 3.5.2
```

```
## Loading required package: lpSolve
```

```
## Warning: package 'lpSolve' was built under R version 3.5.2
```

```
kappa2(data[,c("Class.FRET.PFS", "surrogate_class")])
```

```
##  Cohen's Kappa for 2 Raters (Weights: unweighted)
##
##  Subjects = 182
##    Raters = 2
##     Kappa = 0.177
##
##         z = 2.42
##   p-value = 0.0154
```
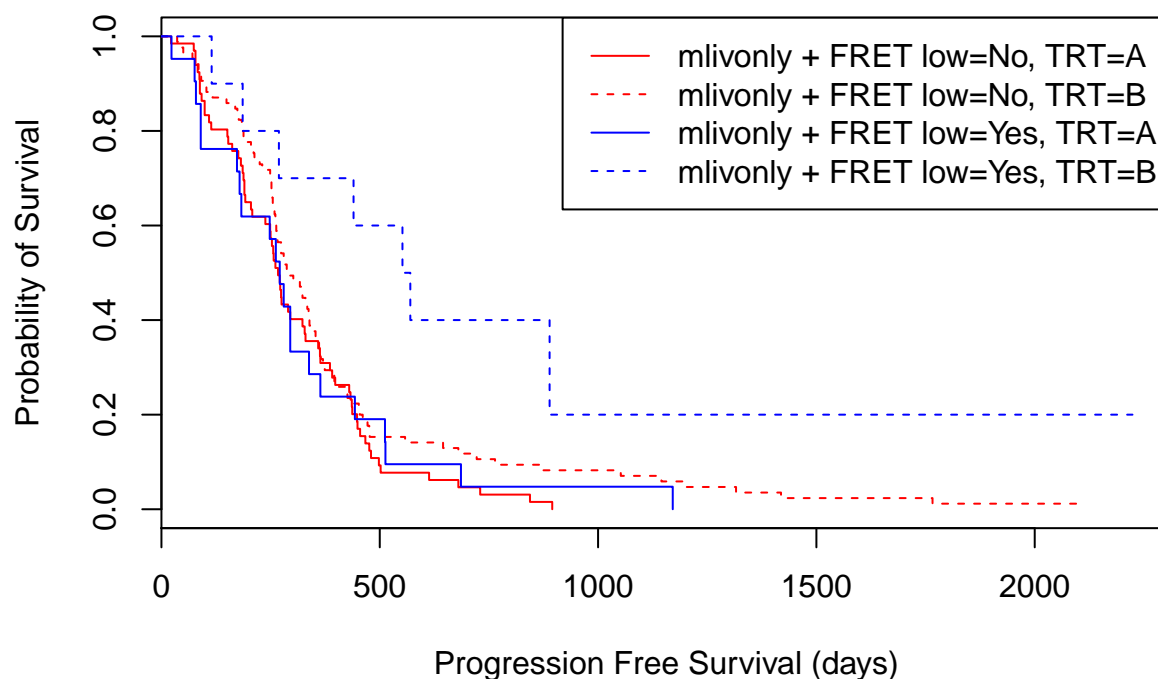
Permutations

Calculate the means and standard deviations from 100,000 random group assignments with the same number per subclass as the real data. Also do permutations test and see how many random group are as, or more, extreme as the real data.

```
## Warning: package 'randomizr' was built under R version 3.5.2
```

```
## [1] "Class number mean and sd."
```

```
##       [,1]         [,2]
## [1,] "3.93±1.67" "19.1±1.67"
## [2,] "27.1±1.67" "132±1.67"
```

```
## [1] "p-values."
```

```
##
##             1         2
##   1 0.0148247 0.0148247
##   2 0.0148247 0.0148247
```

```
## [1] "Permutations:  216 / 10000  => p-value = 0.0216"
```

Figure for paper

```
data$treat  <- ifelse(data$FRETeff=="low" & data$mlivonly=="Yes", "Yes", "No")
```

```
## Call: survfit(formula = SurvObj.pfs ~ treat + TRT, data = data)
##
##    9 observations deleted due to missingness
##                    n events median 0.95LCL 0.95UCL
## treat=No, TRT=A  66     65    267     238     330
## treat=No, TRT=B  85     84    292     263     345
## treat=Yes, TRT=A 21     21    271     179     443
## treat=Yes, TRT=B 10      7    561     269      NA

## Call:
## survdiff(formula = SurvObj.pfs[treat == "Yes"] ~ TRT[treat ==
##     "Yes"], data = data)
##
## n=31, 9 observations deleted due to missingness.
##
##                          N Observed Expected (O-E)^2/E (O-E)^2/V
## TRT[treat == "Yes"]=A 21       21     14.8      2.64      6.01
## TRT[treat == "Yes"]=B 10        7     13.2      2.94      6.01
##
##  Chisq= 6  on 1 degrees of freedom, p= 0.01
```

Session Information

```
sessionInfo()
```

```
## R version 3.5.1 (2018-07-02)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 10 x64 (build 17134)
##
## Matrix products: default
##
## locale:
## [1] LC_COLLATE=English_United Kingdom.1252
## [2] LC_CTYPE=English_United Kingdom.1252
## [3] LC_MONETARY=English_United Kingdom.1252
## [4] LC_NUMERIC=C
## [5] LC_TIME=English_United Kingdom.1252
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
## [1] randomizr_0.16.1 irr_0.84.1       lpSolve_5.6.13   survival_2.42-6
##
## loaded via a namespace (and not attached):
##  [1] Rcpp_1.0.1      lattice_0.20-35 digest_0.6.18   rprojroot_1.3-2
##  [5] grid_3.5.1      backports_1.1.2 magrittr_1.5    evaluate_0.12
##  [9] stringi_1.1.7   Matrix_1.2-14   rmarkdown_1.10  splines_3.5.1
## [13] tools_3.5.1     stringr_1.3.1   yaml_2.2.0      compiler_3.5.1
## [17] htmltools_0.3.6 knitr_1.20
```