

AML blast marker

Rasmus Brøndum

September 11, 2019

Setup workspace

Load libraries

```
library(devtools)
library(survival)
library(xlsx)
library(epitools)
library(GGally)
```

```
## Warning: package 'GGally' was built under R version 4.0.2
```

```
library(Epi)
```

```
## Warning: package 'Epi' was built under R version 4.0.2
```

```
library(knitr)
library(Hmisc)
library(VIM)
```

```
## Warning: package 'VIM' was built under R version 4.0.2
```

```
library(PerformanceAnalytics)
```

```
## Warning: package 'PerformanceAnalytics' was built under R version 4.0.2
```

```
## Warning: package 'xts' was built under R version 4.0.2
```

```
## Warning: package 'zoo' was built under R version 4.0.2
```

```
library(survminer)
```

```
## Warning: package 'survminer' was built under R version 4.0.2
```

```
## Warning: package 'ggpubr' was built under R version 4.0.2
```

```
library(reshape2)
library(ggpmisc)
```

```
## Warning: package 'ggpmisc' was built under R version 4.0.2
```

```
options(knitr.kable.NA = '-')
```

Print Session info

```
session_info()
```

```
## - Session info -----
## setting value
```

```

## version R version 4.0.1 (2020-06-06)
## os      Windows 10 x64
## system  x86_64, mingw32
## ui      RTerm
## language (EN)
## collate Danish_Denmark.1252
## ctype   Danish_Denmark.1252
## tz      Europe/Paris
## date    2020-07-03
##
## - Packages -----
## ! package      * version    date      lib source
## abind          1.4-5      2016-07-21 [1] CRAN (R 4.0.0)
## acepack        1.4.1      2016-10-29 [1] CRAN (R 4.0.0)
## assertthat     0.2.1      2019-03-21 [1] CRAN (R 4.0.0)
## backports      1.1.7      2020-05-13 [1] CRAN (R 4.0.0)
## base64enc      0.1-3      2015-07-28 [1] CRAN (R 4.0.0)
## boot           1.3-25     2020-04-26 [2] CRAN (R 4.0.1)
## broom          0.5.6      2020-04-20 [1] CRAN (R 4.0.0)
## callr          3.4.3      2020-03-28 [1] CRAN (R 4.0.0)
## car            3.0-8      2020-05-21 [1] CRAN (R 4.0.2)
## carData        3.0-4      2020-05-22 [1] CRAN (R 4.0.0)
## cellranger     1.1.0      2016-07-27 [1] CRAN (R 4.0.0)
## checkmate      2.0.0      2020-02-06 [1] CRAN (R 4.0.0)
## class          7.3-17     2020-04-26 [2] CRAN (R 4.0.1)
## cli            2.0.2      2020-02-28 [1] CRAN (R 4.0.0)
## cluster        2.1.0      2019-06-19 [2] CRAN (R 4.0.1)
## cmprsk         2.2-10     2020-06-09 [1] CRAN (R 4.0.2)
## colorspace     * 1.4-1      2019-03-18 [1] CRAN (R 4.0.0)
## crayon         1.3.4      2017-09-16 [1] CRAN (R 4.0.0)
## curl           4.3        2019-12-02 [1] CRAN (R 4.0.0)
## data.table     1.12.8     2019-12-09 [1] CRAN (R 4.0.0)
## DEoptimR       1.0-8      2016-11-19 [1] CRAN (R 4.0.0)
## desc           1.2.0      2018-05-01 [1] CRAN (R 4.0.0)
## devtools       * 2.3.0      2020-04-10 [1] CRAN (R 4.0.0)
## digest         0.6.25     2020-02-23 [1] CRAN (R 4.0.0)
## dplyr          1.0.0      2020-05-29 [1] CRAN (R 4.0.0)
## e1071          1.7-3      2019-11-26 [1] CRAN (R 4.0.2)
## ellipsis       0.3.1      2020-05-15 [1] CRAN (R 4.0.0)
## Epi            * 2.40       2019-11-25 [1] CRAN (R 4.0.2)
## epitools       * 0.5-10.1   2020-03-22 [1] CRAN (R 4.0.0)
## etm            1.1        2020-04-20 [1] CRAN (R 4.0.2)
## evaluate       0.14       2019-05-28 [1] CRAN (R 4.0.0)
## fansi          0.4.1      2020-01-08 [1] CRAN (R 4.0.0)
## forcats        0.5.0      2020-03-01 [1] CRAN (R 4.0.0)
## foreign        0.8-80     2020-05-24 [1] CRAN (R 4.0.0)
## Formula        * 1.2-3      2018-05-03 [1] CRAN (R 4.0.0)
## fs             1.4.1      2020-04-04 [1] CRAN (R 4.0.0)
## generics       0.0.2      2018-11-29 [1] CRAN (R 4.0.0)
## GGally         * 2.0.0      2020-06-06 [1] CRAN (R 4.0.2)
## ggplot2        * 3.3.1      2020-05-28 [1] CRAN (R 4.0.0)
## ggpmisc        * 0.3.5      2020-06-01 [1] CRAN (R 4.0.2)
## ggpubr         * 0.4.0      2020-06-27 [1] CRAN (R 4.0.2)
## ggsignif       0.6.0      2019-08-08 [1] CRAN (R 4.0.2)

```

##	glue	1.4.1	2020-05-13	[1]	CRAN	(R 4.0.0)
##	gridExtra	2.3	2017-09-09	[1]	CRAN	(R 4.0.0)
##	gtable	0.3.0	2019-03-25	[1]	CRAN	(R 4.0.0)
##	haven	2.3.1	2020-06-01	[1]	CRAN	(R 4.0.0)
##	Hmisc	* 4.4-0	2020-03-23	[1]	CRAN	(R 4.0.0)
##	hms	0.5.3	2020-01-08	[1]	CRAN	(R 4.0.0)
##	htmlTable	1.13.3	2019-12-04	[1]	CRAN	(R 4.0.0)
##	htmltools	0.4.0	2019-10-04	[1]	CRAN	(R 4.0.0)
##	htmlwidgets	1.5.1	2019-10-08	[1]	CRAN	(R 4.0.0)
##	jpeg	0.1-8.1	2019-10-24	[1]	CRAN	(R 4.0.0)
##	km.ci	0.5-2	2009-08-30	[1]	CRAN	(R 4.0.2)
##	KMsurv	0.1-5	2012-12-03	[1]	CRAN	(R 4.0.0)
##	knitr	* 1.28	2020-02-06	[1]	CRAN	(R 4.0.0)
##	laeken	0.5.1	2020-02-05	[1]	CRAN	(R 4.0.2)
##	lattice	* 0.20-41	2020-04-02	[2]	CRAN	(R 4.0.1)
##	latticeExtra	0.6-29	2019-12-19	[1]	CRAN	(R 4.0.0)
##	lifecycle	0.2.0	2020-03-06	[1]	CRAN	(R 4.0.0)
##	lmtest	0.9-37	2019-04-30	[1]	CRAN	(R 4.0.2)
##	magrittr	1.5	2014-11-22	[1]	CRAN	(R 4.0.0)
##	MASS	7.3-51.6	2020-04-26	[2]	CRAN	(R 4.0.1)
##	Matrix	1.2-18	2019-11-27	[2]	CRAN	(R 4.0.1)
##	memoise	1.1.0	2017-04-21	[1]	CRAN	(R 4.0.0)
##	mgcv	1.8-31	2019-11-09	[2]	CRAN	(R 4.0.1)
##	munsell	0.5.0	2018-06-12	[1]	CRAN	(R 4.0.0)
##	nlme	3.1-148	2020-05-24	[1]	CRAN	(R 4.0.0)
##	nnet	7.3-14	2020-04-26	[2]	CRAN	(R 4.0.1)
##	numDeriv	2016.8-1.1	2019-06-06	[1]	CRAN	(R 4.0.0)
##	openxlsx	4.1.5	2020-05-06	[1]	CRAN	(R 4.0.2)
##	PerformanceAnalytics	* 2.0.4	2020-02-06	[1]	CRAN	(R 4.0.2)
##	pillar	1.4.4	2020-05-05	[1]	CRAN	(R 4.0.0)
##	pkgbuild	1.0.8	2020-05-07	[1]	CRAN	(R 4.0.0)
##	pkgconfig	2.0.3	2019-09-22	[1]	CRAN	(R 4.0.0)
##	pkgload	1.1.0	2020-05-29	[1]	CRAN	(R 4.0.0)
##	plyr	1.8.6	2020-03-03	[1]	CRAN	(R 4.0.0)
##	png	0.1-7	2013-12-03	[1]	CRAN	(R 4.0.0)
##	prettyunits	1.1.1	2020-01-24	[1]	CRAN	(R 4.0.0)
##	processx	3.4.2	2020-02-09	[1]	CRAN	(R 4.0.0)
##	ps	1.3.3	2020-05-08	[1]	CRAN	(R 4.0.0)
##	purrr	0.3.4	2020-04-17	[1]	CRAN	(R 4.0.0)
##	quadprog	1.5-8	2019-11-20	[1]	CRAN	(R 4.0.0)
##	R6	2.4.1	2019-11-12	[1]	CRAN	(R 4.0.0)
##	ranger	0.12.1	2020-01-10	[1]	CRAN	(R 4.0.2)
##	RColorBrewer	1.1-2	2014-12-07	[1]	CRAN	(R 4.0.0)
##	Rcpp	1.0.4.6	2020-04-09	[1]	CRAN	(R 4.0.0)
##	readxl	1.3.1	2019-03-13	[1]	CRAN	(R 4.0.0)
##	remotes	2.1.1	2020-02-15	[1]	CRAN	(R 4.0.0)
##	reshape	0.8.8	2018-10-23	[1]	CRAN	(R 4.0.2)
##	reshape2	* 1.4.4	2020-04-09	[1]	CRAN	(R 4.0.0)
##	rio	0.5.16	2018-11-26	[1]	CRAN	(R 4.0.2)
##	D rJava	0.9-12	2020-03-24	[1]	CRAN	(R 4.0.0)
##	rlang	0.4.6	2020-05-02	[1]	CRAN	(R 4.0.0)
##	rmarkdown	2.3	2020-06-18	[1]	CRAN	(R 4.0.0)
##	robustbase	0.93-6	2020-03-23	[1]	CRAN	(R 4.0.0)
##	rpart	4.1-15	2019-04-12	[2]	CRAN	(R 4.0.1)

```
##      rprojroot          1.3-2      2018-01-03 [1] CRAN (R 4.0.0)
##      rstatix            0.6.0      2020-06-18 [1] CRAN (R 4.0.2)
##      rstudioapi        0.11       2020-02-07 [1] CRAN (R 4.0.0)
##      scales            1.1.1      2020-05-11 [1] CRAN (R 4.0.0)
##      sessioninfo       1.1.1      2018-11-05 [1] CRAN (R 4.0.0)
##      sp                 1.4-2      2020-05-20 [1] CRAN (R 4.0.2)
##      stringi           1.4.6      2020-02-17 [1] CRAN (R 4.0.0)
##      stringr           1.4.0      2019-02-10 [1] CRAN (R 4.0.0)
##      survival          * 3.1-12    2020-04-10 [2] CRAN (R 4.0.1)
##      survminer         * 0.4.7      2020-05-28 [1] CRAN (R 4.0.2)
##      survMisc          0.5.5      2018-07-05 [1] CRAN (R 4.0.2)
##      testthat          2.3.2      2020-03-02 [1] CRAN (R 4.0.0)
##      tibble            3.0.1      2020-04-20 [1] CRAN (R 4.0.0)
##      tidyr             1.1.0      2020-05-20 [1] CRAN (R 4.0.0)
##      tidyselect        1.1.0      2020-05-11 [1] CRAN (R 4.0.0)
##      usethis           * 1.6.1      2020-04-29 [1] CRAN (R 4.0.0)
##      vcd                1.4-7      2020-04-02 [1] CRAN (R 4.0.2)
##      vctrs             0.3.0      2020-05-11 [1] CRAN (R 4.0.0)
##      VIM               * 6.0.0      2020-05-08 [1] CRAN (R 4.0.2)
##      withr            2.2.0      2020-04-20 [1] CRAN (R 4.0.0)
##      xfun              0.14       2020-05-20 [1] CRAN (R 4.0.0)
##      xlsx              * 0.6.3      2020-02-28 [1] CRAN (R 4.0.0)
##      xlsxjars          0.6.1      2014-08-22 [1] CRAN (R 4.0.0)
##      xtable            1.8-4      2019-04-21 [1] CRAN (R 4.0.0)
##      xts               * 0.12-0     2020-01-19 [1] CRAN (R 4.0.2)
##      yaml              2.2.1      2020-02-01 [1] CRAN (R 4.0.0)
##      zip                2.0.4      2019-09-01 [1] CRAN (R 4.0.2)
##      zoo               * 1.8-8      2020-05-02 [1] CRAN (R 4.0.2)
##
## [1] C:/Users/y12c/Documents/R/win-library/4.0
## [2] C:/Program Files/R/R-4.0.1/library
##
## D -- DLL MD5 mismatch, broken installation.
```

Define custom functions

Function to do uni- and multivariate Poisson regression

```
uniMultiRR <- function(ys, xs, offset = "none", data = data, nDigits = 5, multixs = NA){
  # multixs specifies variables to be included in multivariate analysis
  # if not set, significant from univariate analysis are included
  logitResults <- list()
  for(y in ys){
    ## Do univariate analysis
    for(x in xs){
      if(offset == "none"){
        form <- as.formula(paste(y, "~", x))
      } else{
        form <- as.formula(paste(y, "~", x, "+offset(log(", offset, "))"))
      }
      model <- glm(form, data = data, family = poisson(link = "log"))
      logitResults[[y]][[x]] <- model
    }
  }
  ## Extract coefficients and confidence intervals
```

```

coefs <- lapply(logitResults[[y]], function(x) coef(summary(x)))
confs <- lapply(logitResults[[y]], confint)

RR <- round(sapply(coefs, function(x) exp(x[2,1])), nDigits)
RRc <- round(sapply(confs, function(x) exp(x[2,])), nDigits)
results <- data.frame("RR_uni" = RR,
                      "CI_uni" = paste0(" (",t(RRc)[,1],"-", t(RRc)[,2], ")"),
                      "P_uni" = round(sapply(coefs, function(x) x[2,4]), nDigits))
row.names(results) <- xs

## Pick significant and run multivariate model
if(is.na(multixs[1])){
  sigCoef <- row.names(results[results$P_uni < 0.05,])
} else{
  sigCoef <- multixs
}
if(length(sigCoef) > 0){
  if(offset == "none"){
    form <- as.formula(paste(y, "~", paste(sigCoef, collapse="+")))
  } else{
    form <- as.formula(paste(y, "~", paste(sigCoef, collapse="+"), "+offset(log(", offset,"))"))
  }
  model <- glm(form, data = data, family = poisson(link = "log"))
  logitResults[[y]][["multiVariate"]] <- model

  ## Build output for multivariate model
  RR <- round(exp(coef(model)[-1]), nDigits)
  RRc <- round(exp(confint(model)[-1,, drop = FALSE]), nDigits)
  results[sigCoef, "RR_multi"] <- RR
  results[sigCoef, "CI_multi"] <- paste0(" (",t(RRc[,1]),"-", t(RRc[,2]), ")")
  results[sigCoef, "p_multi"] <- round(coef(summary(model))[-1,4], nDigits)
} else{
  results$RR_multi <- rep(NA, length(xs))
  results$CI_multi <- rep(NA, length(xs))
  results$p_multi <- rep(NA, length(xs))
}

## Save results in list
logitResults[[y]][["resultsTable"]] <- results
}
return(logitResults)
}

```

Function to do uni / multivariate cox analysis possible adjusted for a fixed variable

```

"%w/o%" <- function(x, y) x[!x %in% y] ## x without y
coxR <- function(formula, data, alpha = 0.05, multixs = NA, fix = "Alder"){
  terms <- terms(formula)
  labels <- attr(terms, "term.labels")
  data <- data[!apply(data.frame(data[, labels]), 1, function(x) any(is.na(x))), ]

  for (lab.iter in labels) {
    if (is.factor(data[, lab.iter])) {
      data[,lab.iter] <- droplevels(data[,lab.iter])
    }
  }
}

```

```

}

levels <- vector()
for (lab.iter in labels) {
  if (is.factor(data[,lab.iter])) {
    levels <- c(levels, paste0(lab.iter,levels(data[,lab.iter])))
  } else {
    levels <- c(levels, lab.iter)
  }
}

mat <- matrix(NA, nrow = length(levels), ncol = 6)
rownames(mat) <- levels
colnames(mat) <- rep(c("Hazard ratio", "95% CI", "P-Value"), 2)
mat[, 1] <- 1
y <- rownames(attr(terms,"factors"))[1]
sig.terms <- c()

for (lab.iter in labels) {
  if(is.na(fix[1])){
    formula2 <-
      as.formula((substitute(a ~ b,
                             list(a = as.name(rownames(attr(terms,"factors"))[1]),
                                   b = as.name(lab.iter)))))
  } else{
    formula2 <-
      as.formula(substitute(a ~ b + c,
                             list(a = as.name(rownames(attr(terms,"factors"))[1]),
                                   b = as.name(lab.iter),
                                   c = parse(text = paste(fix, collapse = "+"))[[1]]
                             )
      ))
  }
  cox <- summary(coxph(formula2, data = data))
  COEF <- cox$coefficients

  if(is.na(fix[1])){
    wh <- rownames(COEF)
  } else{
    wh <- grep(lab.iter, rownames(COEF), value = T)
  }

  mat[wh, 1] <- round(COEF[wh,2], 2)
  conf <- round(cox$conf.int[wh, 3:4], 2)

  if (length(wh) > 1) {
    mat[wh, 2] <- apply(conf, 1, function(x)
      paste("(",x[1], "-", x[2], ")", sep = ""))
  }
  if (length(wh) == 1) {
    mat[wh, 2] <- paste("(",conf[1], "-", conf[2], ")", sep = "")
  }
  mat[wh, 3] <- signif(cox$coefficients[wh,5], 2)

```

```

    if(signif(cox$coefficients[wh,5], 2) < alpha) sig.terms <- c(sig.terms, lab.iter)
  }

  if(!is.na(multixs[1])) sig.terms <- multixs

  if(is.na(fix[1])){
    formula2 <- as.formula(paste(as.name(rownames(attr(terms,"factors"))[1]), "~",
                                paste(sig.terms, collapse = " + ")))
  } else{
    formula2 <- as.formula(paste(as.name(rownames(attr(terms,"factors"))[1]), "~",
                                paste(sig.terms, collapse = " + "), "+",
                                paste(fix, collapse = "+")))
  }

  cox <- summary(coxph(formula2, data = data))
  COEF <- cox$coefficients
  if(is.na(fix[1])){
    wh <- rownames(COEF)
  } else{
    wh <- grep(paste(sig.terms, collapse = "|"), row.names(COEF), value = T)
  }
  #for (lab.iter in labels) {
  #  if (is.factor(data[,lab.iter])) {
  #    wh <- gsub(lab.iter, "", wh)
  #  }
  #}
  mat[wh, 4] <- round(COEF[wh,2], 2)
  conf <- round(cox$conf.int[wh, 3:4, drop = FALSE], 2)
  mat[wh, 5] <- apply(conf, 1, function(x)
    paste("(",x[1], "-", x[2], ")", sep = ""))
  mat[wh, 6] <- signif(cox$coefficients[wh,5], 2)
  mat[mat == "NaN"] <- "--"

  list(n = paste("n = ", cox$n, ", number of events = ",
                cox$nevent, sep = ""),
       result = as.data.frame(mat, stringsAsFactors = FALSE))
}

```

Function to split a continous variable by median or quartiles

```

## Function to dichotomize
splitMedian <- function(x){
  y <- ifelse(x > median(x, na.rm = T),"High","Low")
  y <- factor(y, levels = c("Low", "High"))
  return(y)
}

## Function to split in quartiles
splitQuart <- function(x){
  qq <- c(-Inf, quantile(x, c(0.25, 0.50, 0.75), na.rm = T), Inf)
  xx <- cut(x, qq, labels = c("Q1", "Q2", "Q3", "Q4"))
  return(xx)
}

```

Data Management

Load data, remove names from raw data and store a working data file

```
work.data.file <- "../GeneratedData/work.data_DataUpdate120919.Rdata"
if(!file.exists(work.data.file)){
  raw.data <- read.xlsx("../ExternalData/Patients 120919.xlsx",
                        sheetName = "Ark1",
                        encoding = "UTF-8",
                        endRow = 268,
                        check.names = FALSE,
                        password = "xxx") ## Stored in password vault
  dateColumns <- grep("dato", names(raw.data), ignore.case = TRUE)
  work.data <- raw.data[,-dateColumns]
  save(work.data, file = work.data.file)
} else{
  load(work.data.file)
}
```

Check FLT3 vs NPM1 before recoding

```
table(work.data$`FLT3-ITD`, work.data$NPM1, dnn = c("FLT3", "NPM1"))
```

```
##      NPM1
## FLT3   0   1
##      0 157  39
##      1   21  19
```

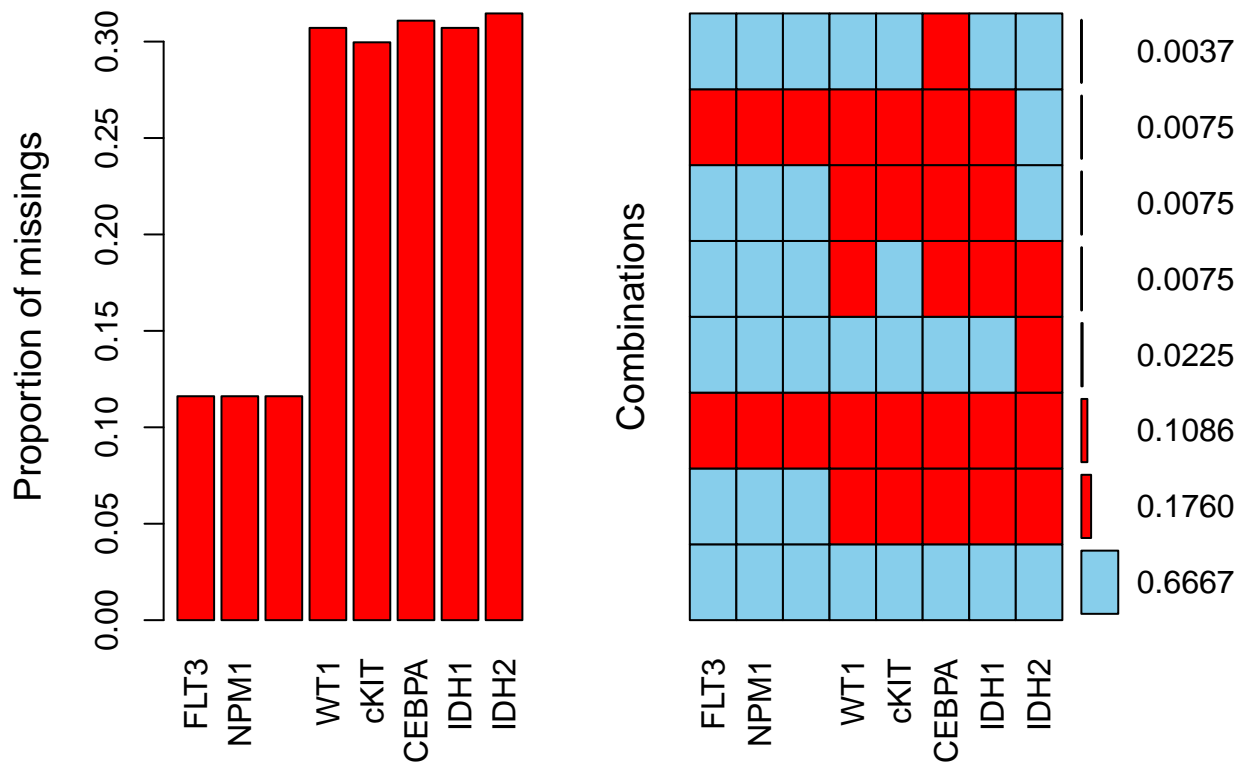
Clean mutation data, Code FLT3+ depending on wt in NPM1+ and vice versa. Check for missing.

```
## FLT3+NPM1+ requires mutation in both
work.data$FLT3_NPM1 <- as.numeric(work.data$`FLT3-ITD` + work.data$NPM1 == 2)

## FLT3 positive requires mutation in FLT3-ITD and wildtype in NPM1
work.data$FLT3 <- as.numeric(work.data$`FLT3-ITD` == 1 & work.data$NPM1 == 0)

## NPM1 positive requires wildtype in FLT3-ITD and mutation in NPM1
work.data$NPM1 <- as.numeric(work.data$`FLT3-ITD` == 0 & work.data$NPM1 == 1)

mutations <- c("FLT3", "NPM1", "FLT3_NPM1", "WT1", "cKIT", "CEBPA", "IDH1", "IDH2")
aggr(work.data[,mutations], numbers = T)
```

Cleanup names of flow data. All selected markers are CD45low, SSClow and CD14neg, so this part of the name is deleted.

```
counts <- grep("Count", names(work.data))
names(work.data)[counts] <- gsub("\\+", "pos", names(work.data)[counts])
names(work.data)[counts] <- gsub("\\-", "neg", names(work.data)[counts])
names(work.data)[counts] <- gsub(" \\| Count", "", names(work.data)[counts])
names(work.data)[counts] <- gsub("/", ".", names(work.data)[counts])
flows <- c("CD34pos",
           "CD117pos",
           "CD123pos",
           "CLEC12Apos",
           "CD34pos.CD123pos",
           "CD34neg.CD123pos",
           "CD34pos.CD117pos",
           "CD34neg.CD117pos",
           "CD34pos.CLEC12Apos",
           "CD34neg.CLEC12Apos")
```

flows

```
## [1] "CD34pos"          "CD117pos"         "CD123pos"
## [4] "CLEC12Apos"       "CD34pos.CD123pos" "CD34neg.CD123pos"
## [7] "CD34pos.CD117pos" "CD34neg.CD117pos" "CD34pos.CLEC12Apos"
## [10] "CD34neg.CLEC12Apos"
```

Generate relative counts, normalized by number of live counts

```
flow2 <- (work.data[, flows] / work.data[, "Live cells"]) * 100
names(flow2) <- paste0(flows, ".frq")
work.data <- cbind(work.data, flow2)
flows <- paste0(flows, ".frq")
```

Include leukocyttter in analysis

```
work.data$leukocyttter <- work.data$`Leukocyttter (mia/l)`
flows <- c(flows, "leukocyttter")
```

Descriptive statistics

Tabulate mutations and MRC

```
muts <- rbind(table(work.data$FLT3),
               table(work.data$NPM1),
               table(work.data$FLT3_NPM1),
               table(work.data$WT1),
               table(work.data$cKIT),
               table(work.data$CEBPA),
               table(work.data$IDH1),
               table(work.data$IDH2))
row.names(muts) <- c("FLT3", "NPM1", "FLT3+NPM1", "WT1", "cKIT", "CEBPA", "IDH1", "IDH2")
colnames(muts) <- c("WildType", "Mutation")
kable(muts,
      caption = "Mutation or wildtype for tested genes across patients")
```

Table 1: Mutation or wildtype for tested genes across patients

	WildType	Mutation
FLT3	215	21
NPM1	197	39
FLT3+NPM1	217	19
WT1	181	4
cKIT	184	3
CEBPA	173	11
IDH1	168	17
IDH2	156	27

Summarize MRC

```
kable(table(work.data$`MRC classification`),
      caption = "Summary of MRC classification")
```

Table 2: Summary of MRC classification

Var1	Freq
1	20
2	187
3	54

Correlation of flow markers

```
flowCor <- rcorr(as.matrix(work.data[,flows]))
corM <- matrix(0, ncol = ncol(flowCor$P), nrow = nrow(flowCor$P))
corM[upper.tri(corM)] <- flowCor$P[upper.tri(flowCor$P)]
corM[lower.tri(corM)] <- flowCor$r[lower.tri(flowCor$r)]
row.names(corM) <- sub(".frq", "", row.names(flowCor$P))

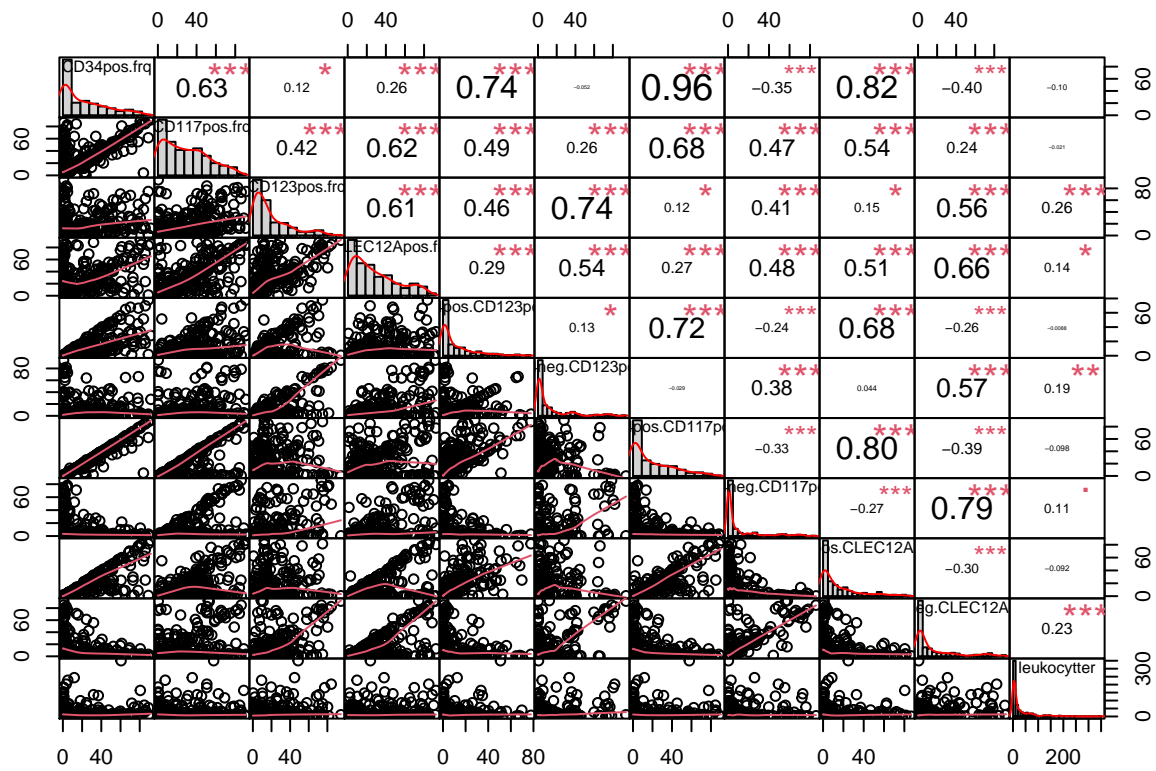
kable(corM,
      col.names = 1:ncol(corM),
      caption = paste("Pairwise correlation of Flow markers.",
                      "Correlation on lower diagonal, p-value",
                      " for pairwise test on upper diagonal"),
      digits = 2)
```

Table 3: Pairwise correlation of Flow markers. Correlation on lower diagonal, p-value for pairwise test on upper diagonal

	1	2	3	4	5	6	7	8	9	10	11
CD34pos	0.00	0.00	0.04	0.00	0.00	0.40	0.00	0.00	0.00	0.00	0.10
CD117pos	0.63	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.74
CD123pos	0.12	0.42	0.00	0.00	0.00	0.00	0.05	0.00	0.02	0.00	0.00
CLEC12Apos	0.26	0.62	0.61	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03
CD34pos.CD123pos	0.74	0.49	0.46	0.29	0.00	0.03	0.00	0.00	0.00	0.00	0.89
CD34neg.CD123pos	-0.05	0.26	0.74	0.54	0.13	0.00	0.63	0.00	0.47	0.00	0.00
CD34pos.CD117pos	0.96	0.68	0.12	0.27	0.72	-0.03	0.00	0.00	0.00	0.00	0.11
CD34neg.CD117pos	-0.35	0.47	0.41	0.48	-0.24	0.38	-0.33	0.00	0.00	0.00	0.07
CD34pos.CLEC12Apos	0.82	0.54	0.15	0.51	0.68	0.04	0.80	-0.27	0.00	0.00	0.13
CD34neg.CLEC12Apos	-0.40	0.24	0.56	0.66	-0.26	0.57	-0.39	0.79	-0.30	0.00	0.00
leukocyttter	-0.10	-0.02	0.26	0.14	-0.01	0.19	-0.10	0.11	-0.09	0.23	0.00

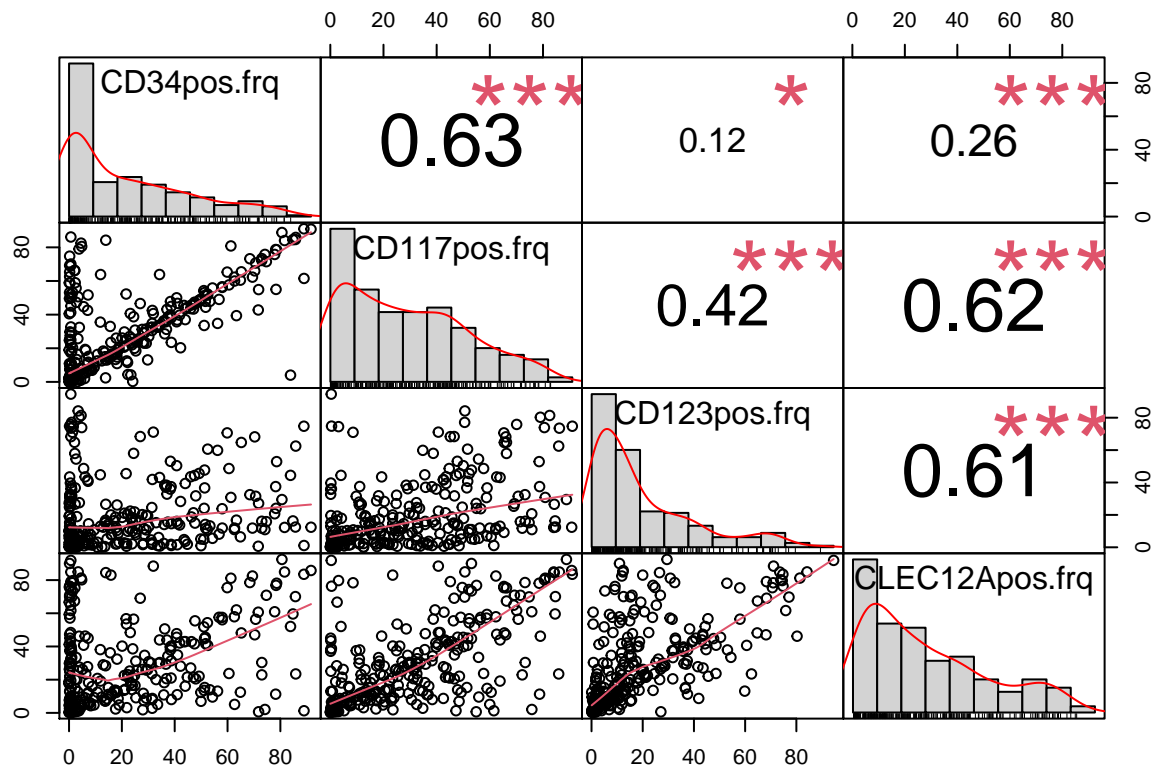
Plot entire correlation matrix

```
chart.Correlation(work.data[,flows], histogram=TRUE, pch=10)
```



Plot Just the four markers in the manuscript

```
chart.Correlation(work.data[,flows[1:4]], histogram=TRUE, pch=19, method = c("pearson"))
```



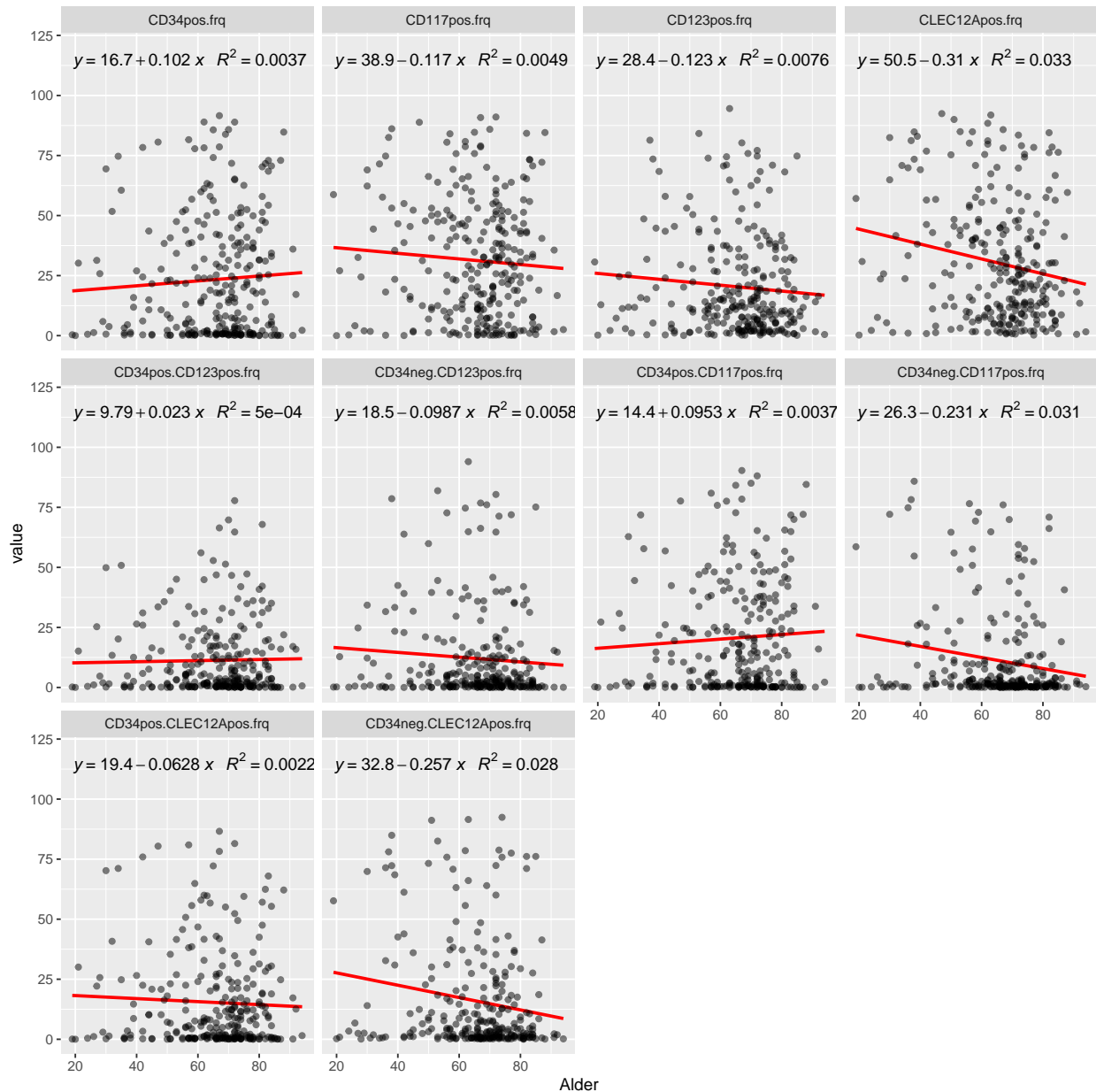
Regress age on flow markers

```
tempData <- work.data[,c("ID", "Alder", flows[-length(flows)])]
tempData <- melt(tempData, id.vars = c("ID", "Alder"))
ggplot(tempData, aes(x = Alder, y = value)) +
  facet_wrap(~variable) +
  geom_smooth(method = "lm", se=FALSE, color="red", formula = y ~ x) +
  ylim(c(0,120)) +
  stat_poly_eq(formula = y ~ x,
    aes(label = paste(..eq.label.., ..rr.label.., sep = "~~~")),
    parse = TRUE) +
  geom_point(alpha = 0.5)
```

Warning: Removed 10 rows containing non-finite values (stat_smooth).

Warning: Removed 10 rows containing non-finite values (stat_poly_eq).

Warning: Removed 10 rows containing missing values (geom_point).



Tables with tests of dichotomized immunophenotypes

MRC vs flow markers.

```
tempData <- data.frame(lapply(subset(work.data, select = c(flows)), splitMedian),
                        "MRC"=work.data$`MRC classification`)
temp <- list()
for(flow in flows){
  tab <- table(tempData[, flow], tempData[, "MRC"])
  tab.test <- fisher.test(tab)

  tab2 <- cbind(matrix(tab, ncol = 3), c(tab.test$p.value, NA))
  row.names(tab2) <- paste(flow, row.names(tab), sep = ".")
}
```

```

colnames(tab2) <- c(colnames(tab), "p-value")

temp[[flow]] <- tab2
}

```

Table 4: MRC classification vs above/below median values for flow markers. P-value from fishers exact test

	1	2	3	p-value
CD34pos.frq.Low	4	107	20	0.001
CD34pos.frq.High	16	80	34	-
CD117pos.frq.Low	6	92	30	0.161
CD117pos.frq.High	14	95	24	-
CD123pos.frq.Low	7	89	34	0.056
CD123pos.frq.High	13	98	20	-
CLEC12Apos.frq.Low	3	88	39	0.000
CLEC12Apos.frq.High	17	99	15	-
CD34pos.CD123pos.frq.Low	2	106	23	0.000
CD34pos.CD123pos.frq.High	18	81	31	-
CD34neg.CD123pos.frq.Low	8	89	34	0.090
CD34neg.CD123pos.frq.High	12	98	20	-
CD34pos.CD117pos.frq.Low	3	106	22	0.000
CD34pos.CD117pos.frq.High	17	81	32	-
CD34neg.CD117pos.frq.Low	14	82	32	0.021
CD34neg.CD117pos.frq.High	6	105	22	-
CD34pos.CLEC12Apos.frq.Low	3	104	24	0.001
CD34pos.CLEC12Apos.frq.High	17	83	30	-
CD34neg.CLEC12Apos.frq.Low	14	81	34	0.006
CD34neg.CLEC12Apos.frq.High	6	106	20	-
leukocytter.Low	7	89	34	0.063
leukocytter.High	13	96	20	-

Mutations vs flow markers

Check mutations vs flow markers. This corresponds to just a single point for the ROC analysis done later

```

tempData <- data.frame(lapply(subset(work.data, select = c(flows)), splitMedian))
tempData <- cbind(tempData, work.data[,mutations])
temp <- list()
for(mut in mutations){
  for(flow in flows){
    tab <- table(tempData[, flow], tempData[,mut])
    tab.test <- fisher.test(tab)

    tab2 <- matrix(tab, ncol = 2)
    percent <- round(tab2[,2] *100 / rowSums(tab2),1)
    tab2 <- cbind(tab2, percent, c(tab.test$p.value, NA))

    row.names(tab2) <- paste(flow,row.names(tab), sep = ".")
    colnames(tab2) <- c("wildtype", "mutation", "percent mut","p-value")

    temp[[mut]][[flow]] <- tab2
  }
}

```

}

Table 5: Percent patients with mutation vs above/below median values for flow markers for FLT3. P-value from fishers exact test.

	wildtype	mutation	percent mut	p-value
CD34pos.frq.Low	106	9	7.8	0.651
CD34pos.frq.High	109	12	9.9	-
CD117pos.frq.Low	107	9	7.8	0.650
CD117pos.frq.High	108	12	10.0	-
CD123pos.frq.Low	107	5	4.5	0.037
CD123pos.frq.High	108	16	12.9	-
CLEC12Apos.frq.Low	110	5	4.3	0.021
CLEC12Apos.frq.High	105	16	13.2	-
CD34pos.CD123pos.frq.Low	108	7	6.1	0.172
CD34pos.CD123pos.frq.High	107	14	11.6	-
CD34neg.CD123pos.frq.Low	114	4	3.4	0.005
CD34neg.CD123pos.frq.High	101	17	14.4	-
CD34pos.CD117pos.frq.Low	106	9	7.8	0.651
CD34pos.CD117pos.frq.High	109	12	9.9	-
CD34neg.CD117pos.frq.Low	113	10	8.1	0.820
CD34neg.CD117pos.frq.High	102	11	9.7	-
CD34pos.CLEC12Apos.frq.Low	107	9	7.8	0.650
CD34pos.CLEC12Apos.frq.High	108	12	10.0	-
CD34neg.CLEC12Apos.frq.Low	110	8	6.8	0.361
CD34neg.CLEC12Apos.frq.High	105	13	11.0	-
leukocytter.Low	106	8	7.0	0.364
leukocytter.High	107	13	10.8	-

Table 6: Percent patients with mutation vs above/below median values for flow markers for NPM1. P-value from fishers exact test.

	wildtype	mutation	percent mut	p-value
CD34pos.frq.Low	78	37	32.2	0.000
CD34pos.frq.High	119	2	1.7	-
CD117pos.frq.Low	88	28	24.1	0.003
CD117pos.frq.High	109	11	9.2	-
CD123pos.frq.Low	94	18	16.1	1.000
CD123pos.frq.High	103	21	16.9	-
CLEC12Apos.frq.Low	100	15	13.0	0.167
CLEC12Apos.frq.High	97	24	19.8	-
CD34pos.CD123pos.frq.Low	82	33	28.7	0.000
CD34pos.CD123pos.frq.High	115	6	5.0	-
CD34neg.CD123pos.frq.Low	97	21	17.8	0.726
CD34neg.CD123pos.frq.High	100	18	15.3	-
CD34pos.CD117pos.frq.Low	79	36	31.3	0.000
CD34pos.CD117pos.frq.High	118	3	2.5	-
CD34neg.CD117pos.frq.Low	112	11	8.9	0.001
CD34neg.CD117pos.frq.High	85	28	24.8	-
CD34pos.CLEC12Apos.frq.Low	83	33	28.4	0.000
CD34pos.CLEC12Apos.frq.High	114	6	5.0	-
CD34neg.CLEC12Apos.frq.Low	113	5	4.2	0.000

	wildtype	mutation	percent mut	p-value
CD34neg.CLEC12Apos.frq.High	84	34	28.8	-
leukocytter.Low	100	14	12.3	0.115
leukocytter.High	96	24	20.0	-

Table 7: Percent patients with mutation vs above/below median values for flow markers for FLT3_NPM1. P-value from fishers exact test.

	wildtype	mutation	percent mut	p-value
CD34pos.frq.Low	98	17	14.8	0.000
CD34pos.frq.High	119	2	1.7	-
CD117pos.frq.Low	109	7	6.0	0.340
CD117pos.frq.High	108	12	10.0	-
CD123pos.frq.Low	109	3	2.7	0.004
CD123pos.frq.High	108	16	12.9	-
CLEC12Apos.frq.Low	111	4	3.5	0.015
CLEC12Apos.frq.High	106	15	12.4	-
CD34pos.CD123pos.frq.Low	103	12	10.4	0.234
CD34pos.CD123pos.frq.High	114	7	5.8	-
CD34neg.CD123pos.frq.Low	114	4	3.4	0.015
CD34neg.CD123pos.frq.High	103	15	12.7	-
CD34pos.CD117pos.frq.Low	98	17	14.8	0.000
CD34pos.CD117pos.frq.High	119	2	1.7	-
CD34neg.CD117pos.frq.Low	123	0	0.0	0.000
CD34neg.CD117pos.frq.High	94	19	16.8	-
CD34pos.CLEC12Apos.frq.Low	100	16	13.8	0.001
CD34pos.CLEC12Apos.frq.High	117	3	2.5	-
CD34neg.CLEC12Apos.frq.Low	116	2	1.7	0.000
CD34neg.CLEC12Apos.frq.High	101	17	14.4	-
leukocytter.Low	112	2	1.8	0.000
leukocytter.High	103	17	14.2	-

Table 8: Percent patients with mutation vs above/below median values for flow markers for WT1. P-value from fishers exact test.

	wildtype	mutation	percent mut	p-value
CD34pos.frq.Low	85	2	2.3	1.000
CD34pos.frq.High	96	2	2.0	-
CD117pos.frq.Low	84	2	2.3	1.000
CD117pos.frq.High	97	2	2.0	-
CD123pos.frq.Low	88	2	2.2	1.000
CD123pos.frq.High	93	2	2.1	-
CLEC12Apos.frq.Low	88	2	2.2	1.000
CLEC12Apos.frq.High	93	2	2.1	-
CD34pos.CD123pos.frq.Low	86	2	2.3	1.000
CD34pos.CD123pos.frq.High	95	2	2.1	-
CD34neg.CD123pos.frq.Low	90	1	1.1	0.621
CD34neg.CD123pos.frq.High	91	3	3.2	-
CD34pos.CD117pos.frq.Low	85	2	2.3	1.000
CD34pos.CD117pos.frq.High	96	2	2.0	-

	wildtype	mutation	percent mut	p-value
CD34neg.CD117pos.frq.Low	95	2	2.1	1.000
CD34neg.CD117pos.frq.High	86	2	2.3	-
CD34pos.CLEC12Apos.frq.Low	87	2	2.2	1.000
CD34pos.CLEC12Apos.frq.High	94	2	2.1	-
CD34neg.CLEC12Apos.frq.Low	88	2	2.2	1.000
CD34neg.CLEC12Apos.frq.High	93	2	2.1	-
leukocytter.Low	87	2	2.2	1.000
leukocytter.High	92	2	2.1	-

Table 9: Percent patients with mutation vs above/below median values for flow markers for cKIT. P-value from fishers exact test.

	wildtype	mutation	percent mut	p-value
CD34pos.frq.Low	87	0	0.0	0.250
CD34pos.frq.High	97	3	3.0	-
CD117pos.frq.Low	85	2	2.3	0.598
CD117pos.frq.High	99	1	1.0	-
CD123pos.frq.Low	91	0	0.0	0.247
CD123pos.frq.High	93	3	3.1	-
CLEC12Apos.frq.Low	90	0	0.0	0.247
CLEC12Apos.frq.High	94	3	3.1	-
CD34pos.CD123pos.frq.Low	88	0	0.0	0.249
CD34pos.CD123pos.frq.High	96	3	3.0	-
CD34neg.CD123pos.frq.Low	91	0	0.0	0.247
CD34neg.CD123pos.frq.High	93	3	3.1	-
CD34pos.CD117pos.frq.Low	87	0	0.0	0.250
CD34pos.CD117pos.frq.High	97	3	3.0	-
CD34neg.CD117pos.frq.Low	97	2	2.0	1.000
CD34neg.CD117pos.frq.High	87	1	1.1	-
CD34pos.CLEC12Apos.frq.Low	89	0	0.0	0.248
CD34pos.CLEC12Apos.frq.High	95	3	3.1	-
CD34neg.CLEC12Apos.frq.Low	90	2	2.2	0.617
CD34neg.CLEC12Apos.frq.High	94	1	1.1	-
leukocytter.Low	89	0	0.0	0.247
leukocytter.High	93	3	3.1	-

Table 10: Percent patients with mutation vs above/below median values for flow markers for CEBPA. P-value from fishers exact test.

	wildtype	mutation	percent mut	p-value
CD34pos.frq.Low	85	1	1.2	0.011
CD34pos.frq.High	88	10	10.2	-
CD117pos.frq.Low	82	3	3.5	0.228
CD117pos.frq.High	91	8	8.1	-
CD123pos.frq.Low	85	4	4.5	0.539
CD123pos.frq.High	88	7	7.4	-
CLEC12Apos.frq.Low	88	1	1.1	0.010
CLEC12Apos.frq.High	85	10	10.5	-
CD34pos.CD123pos.frq.Low	84	3	3.4	0.220
CD34pos.CD123pos.frq.High	89	8	8.2	-

	wildtype	mutation	percent mut	p-value
CD34neg.CD123pos.frq.Low	88	2	2.2	0.058
CD34neg.CD123pos.frq.High	85	9	9.6	-
CD34pos.CD117pos.frq.Low	85	1	1.2	0.011
CD34pos.CD117pos.frq.High	88	10	10.2	-
CD34neg.CD117pos.frq.Low	91	5	5.2	0.760
CD34neg.CD117pos.frq.High	82	6	6.8	-
CD34pos.CLEC12Apos.frq.Low	87	1	1.1	0.010
CD34pos.CLEC12Apos.frq.High	86	10	10.4	-
CD34neg.CLEC12Apos.frq.Low	86	4	4.4	0.537
CD34neg.CLEC12Apos.frq.High	87	7	7.4	-
leukocytter.Low	83	5	5.7	1.000
leukocytter.High	88	6	6.4	-

Table 11: Percent patients with mutation vs above/below median values for flow markers for IDH1. P-value from fishers exact test.

	wildtype	mutation	percent mut	p-value
CD34pos.frq.Low	82	5	5.7	0.201
CD34pos.frq.High	86	12	12.2	-
CD117pos.frq.Low	80	6	7.0	0.446
CD117pos.frq.High	88	11	11.1	-
CD123pos.frq.Low	84	6	6.7	0.312
CD123pos.frq.High	84	11	11.6	-
CLEC12Apos.frq.Low	82	8	8.9	1.000
CLEC12Apos.frq.High	86	9	9.5	-
CD34pos.CD123pos.frq.Low	82	6	6.8	0.319
CD34pos.CD123pos.frq.High	86	11	11.3	-
CD34neg.CD123pos.frq.Low	82	9	9.9	0.803
CD34neg.CD123pos.frq.High	86	8	8.5	-
CD34pos.CD117pos.frq.Low	81	6	6.9	0.445
CD34pos.CD117pos.frq.High	87	11	11.2	-
CD34neg.CD117pos.frq.Low	88	9	9.3	1.000
CD34neg.CD117pos.frq.High	80	8	9.1	-
CD34pos.CLEC12Apos.frq.Low	84	5	5.6	0.130
CD34pos.CLEC12Apos.frq.High	84	12	12.5	-
CD34neg.CLEC12Apos.frq.Low	79	11	12.2	0.206
CD34neg.CLEC12Apos.frq.High	89	6	6.3	-
leukocytter.Low	76	13	14.6	0.021
leukocytter.High	90	4	4.3	-

Table 12: Percent patients with mutation vs above/below median values for flow markers for IDH2. P-value from fishers exact test.

	wildtype	mutation	percent mut	p-value
CD34pos.frq.Low	76	12	13.6	0.835
CD34pos.frq.High	80	15	15.8	-
CD117pos.frq.Low	74	13	14.9	1.000
CD117pos.frq.High	82	14	14.6	-
CD123pos.frq.Low	75	15	16.7	0.535
CD123pos.frq.High	81	12	12.9	-

	wildtype	mutation	percent mut	p-value
CLEC12Apos.frq.Low	78	11	12.4	0.410
CLEC12Apos.frq.High	78	16	17.0	-
CD34pos.CD123pos.frq.Low	73	15	17.0	0.413
CD34pos.CD123pos.frq.High	83	12	12.6	-
CD34neg.CD123pos.frq.Low	72	17	19.1	0.144
CD34neg.CD123pos.frq.High	84	10	10.6	-
CD34pos.CD117pos.frq.Low	77	11	12.5	0.532
CD34pos.CD117pos.frq.High	79	16	16.8	-
CD34neg.CD117pos.frq.Low	84	12	12.5	0.408
CD34neg.CD117pos.frq.High	72	15	17.2	-
CD34pos.CLEC12Apos.frq.Low	76	11	12.6	0.533
CD34pos.CLEC12Apos.frq.High	80	16	16.7	-
CD34neg.CLEC12Apos.frq.Low	74	14	15.9	0.683
CD34neg.CLEC12Apos.frq.High	82	13	13.7	-
leukocyttter.Low	73	17	18.9	0.094
leukocyttter.High	82	9	9.9	-

Binary regression: Mutations or risk score vs flow markers

Recode MRC classification as binary variables with one group vs others.

```
work.data$MRC.Favorable <- ifelse(work.data$`MRC classification` == 1, 1, 0)
work.data$MRC.Intermediate <- ifelse(work.data$`MRC classification` == 2, 1, 0)
work.data$MRC.Adverse <- ifelse(work.data$`MRC classification` == 3, 1, 0)
risks <- c("MRC.Favorable", "MRC.Intermediate", "MRC.Adverse")
```

Immunophenotypes vs mutational aberrations

```
mutFlowResults <- uniMultiRR(ys = mutations,
                             xs = flows,
                             data = work.data,
                             nDigits = 4,
                             multixs = flows[1:4])
```

Table 13: Poisson regression result for FLT3

	Risk Ratio	95% CI	P-Value	Risk Ratio	95% CI	P-Value
CD34pos.frq	1.0047	(0.9877-1.0202)	0.5650	1.0027	(0.9843-1.0214)	0.7695
CD117pos.frq	1.0112	(0.9946-1.0277)	0.1786	0.9937	(0.9727-1.0155)	0.5601
CD123pos.frq	1.0194	(1.0024-1.0354)	0.0183	1.0053	(0.9847-1.0264)	0.6108
CLEC12Apos.frq	1.0237	(1.0083-1.0397)	0.0025	1.0237	(1.0015-1.0463)	0.0352
CD34pos.CD123pos.frq	1.0094	(0.9825-1.0316)	0.4471	-	-	-
CD34neg.CD123pos.frq	1.0188	(1.002-1.0335)	0.0163	-	-	-
CD34pos.CD117pos.frq	1.0051	(0.9869-1.0215)	0.5562	-	-	-
CD34neg.CD117pos.frq	1.0098	(0.9899-1.0263)	0.2784	-	-	-
CD34pos.CLEC12Apos.frq	1.0172	(0.9993-1.0333)	0.0434	-	-	-
CD34neg.CLEC12Apos.frq	1.0115	(0.9954-1.0258)	0.1286	-	-	-
leukocyttter	1.0059	(0.9992-1.0108)	0.0401	-	-	-

Table 14: Poisson regression result for NPM1

	Risk Ratio	95% CI	P-Value	Risk Ratio	95% CI	P-Value
CD34pos.frq	0.9221	(0.8792-0.9544)	0.0001	0.9300	(0.888-0.9616)	0.0003
CD117pos.frq	0.9829	(0.9681-0.9964)	0.0184	0.9859	(0.9688-1.002)	0.0938
CD123pos.frq	1.0076	(0.9937-1.0203)	0.2586	0.9913	(0.9738-1.0094)	0.3365
CLEC12Apos.frq	1.0096	(0.9979-1.0211)	0.0993	1.0247	(1.0052-1.0435)	0.0102
CD34pos.CD123pos.frq	0.9196	(0.8637-0.9627)	0.0022	-	-	-
CD34neg.CD123pos.frq	1.0103	(0.9963-1.0224)	0.1180	-	-	-
CD34pos.CD117pos.frq	0.9138	(0.8642-0.9507)	0.0002	-	-	-
CD34neg.CD117pos.frq	1.0155	(1.0027-1.027)	0.0110	-	-	-
CD34pos.CLEC12Apos.frq	0.9285	(0.8807-0.9648)	0.0013	-	-	-
CD34neg.CLEC12Apos.frq	1.0224	(1.0123-1.0322)	0.0000	-	-	-
leukocytter	1.0055	(1.0006-1.0094)	0.0120	-	-	-

Table 15: Poisson regression result for FLT3_NPM1

	Risk Ratio	95% CI	P-Value	Risk Ratio	95% CI	P-Value
CD34pos.frq	0.9467	(0.9008-0.9792)	0.0084	0.9463	(0.9003-0.9782)	0.0068
CD117pos.frq	1.0055	(0.9877-1.0228)	0.5358	1.0094	(0.9898-1.0315)	0.3672
CD123pos.frq	1.0292	(1.0125-1.0458)	0.0004	1.0399	(1.011-1.0749)	0.0123
CLEC12Apos.frq	1.0121	(0.9953-1.0284)	0.1451	0.9774	(0.9463-1.0063)	0.1427
CD34pos.CD123pos.frq	0.9682	(0.9175-1.0049)	0.1550	-	-	-
CD34neg.CD123pos.frq	1.0261	(1.0102-1.0407)	0.0006	-	-	-
CD34pos.CD117pos.frq	0.9481	(0.9-0.9818)	0.0136	-	-	-
CD34neg.CD117pos.frq	1.0297	(1.0141-1.0447)	0.0001	-	-	-
CD34pos.CLEC12Apos.frq	0.9377	(0.8724-0.982)	0.0295	-	-	-
CD34neg.CLEC12Apos.frq	1.0265	(1.0123-1.0405)	0.0002	-	-	-
leukocytter	1.0070	(1.0007-1.0117)	0.0103	-	-	-

Table 16: Poisson regression result for WT1

	Risk Ratio	95% CI	P-Value	Risk Ratio	95% CI	P-Value
CD34pos.frq	0.9936	(0.9441-1.03)	0.7577	0.9940	(0.9405-1.0525)	0.8169
CD117pos.frq	0.9990	(0.9571-1.0369)	0.9604	0.9995	(0.9393-1.0568)	0.9852
CD123pos.frq	1.0140	(0.9712-1.0503)	0.4615	1.0261	(0.9694-1.0941)	0.3904
CLEC12Apos.frq	1.0005	(0.9586-1.0361)	0.9781	0.9845	(0.9202-1.0391)	0.6043
CD34pos.CD123pos.frq	1.0134	(0.949-1.0578)	0.6032	-	-	-
CD34neg.CD123pos.frq	0.9736	(0.8469-1.026)	0.5190	-	-	-
CD34pos.CD117pos.frq	0.9965	(0.9457-1.0335)	0.8685	-	-	-
CD34neg.CD117pos.frq	1.0037	(0.939-1.0422)	0.8748	-	-	-
CD34pos.CLEC12Apos.frq	1.0044	(0.9488-1.0424)	0.8439	-	-	-
CD34neg.CLEC12Apos.frq	0.9955	(0.9333-1.0316)	0.8414	-	-	-
leukocytter	0.9982	(0.963-1.0124)	0.8748	-	-	-

Table 17: Poisson regression result for cKIT

	Risk Ratio	95% CI	P-Value	Risk Ratio	95% CI	P-Value
CD34pos.frq	1.0061	(0.958-1.0464)	0.7712	1.0152	(0.9567-1.1033)	0.6569
CD117pos.frq	1.0008	(0.952-1.0451)	0.9721	0.9766	(0.8923-1.0485)	0.5401
CD123pos.frq	1.0038	(0.9437-1.0472)	0.8772	0.9930	(0.9251-1.0521)	0.8176
CLEC12Apos.frq	1.0137	(0.9703-1.0564)	0.5032	1.0252	(0.9663-1.0865)	0.3965
CD34pos.CD123pos.frq	1.0151	(0.9395-1.0654)	0.6072	-	-	-
CD34neg.CD123pos.frq	1.0129	(0.9572-1.0506)	0.5452	-	-	-
CD34pos.CD117pos.frq	1.0015	(0.946-1.0431)	0.9468	-	-	-
CD34neg.CD117pos.frq	0.9999	(0.9115-1.0454)	0.9962	-	-	-
CD34pos.CLEC12Apos.frq	1.0216	(0.9727-1.0628)	0.3047	-	-	-
CD34neg.CLEC12Apos.frq	0.9909	(0.9044-1.0342)	0.7533	-	-	-
leukocytter	0.9988	(0.9571-1.014)	0.9221	-	-	-

Table 18: Poisson regression result for CEBPA

	Risk Ratio	95% CI	P-Value	Risk Ratio	95% CI	P-Value
CD34pos.frq	1.0274	(1.007-1.0492)	0.0084	1.0224	(0.9947-1.061)	0.1685
CD117pos.frq	1.0269	(1.0042-1.0518)	0.0224	0.9798	(0.9398-1.0193)	0.3153
CD123pos.frq	1.0020	(0.9736-1.0255)	0.8780	0.9801	(0.9472-1.0079)	0.1917
CLEC12Apos.frq	1.0392	(1.0168-1.066)	0.0011	1.0511	(1.0195-1.0866)	0.0020
CD34pos.CD123pos.frq	1.0359	(1.0094-1.06)	0.0039	-	-	-
CD34neg.CD123pos.frq	1.0112	(0.9855-1.0317)	0.3237	-	-	-
CD34pos.CD117pos.frq	1.0304	(1.0098-1.0522)	0.0037	-	-	-
CD34neg.CD117pos.frq	0.9807	(0.9206-1.015)	0.3966	-	-	-
CD34pos.CLEC12Apos.frq	1.0417	(1.0209-1.064)	0.0001	-	-	-
CD34neg.CLEC12Apos.frq	0.9992	(0.9695-1.0211)	0.9471	-	-	-
leukocytter	0.9988	(0.9818-1.0084)	0.8558	-	-	-

Table 19: Poisson regression result for IDH1

	Risk Ratio	95% CI	P-Value	Risk Ratio	95% CI	P-Value
CD34pos.frq	1.0098	(0.9923-1.0264)	0.2493	1.0144	(0.9907-1.0438)	0.2703
CD117pos.frq	1.0078	(0.9894-1.0259)	0.3977	0.9920	(0.9608-1.0208)	0.5927
CD123pos.frq	1.0193	(1.001-1.0366)	0.0301	1.0316	(1.0064-1.059)	0.0155
CLEC12Apos.frq	1.0037	(0.9854-1.0208)	0.6766	0.9860	(0.96-1.01)	0.2709
CD34pos.CD123pos.frq	1.0358	(1.0148-1.0552)	0.0003	-	-	-
CD34neg.CD123pos.frq	1.0073	(0.9854-1.0248)	0.4586	-	-	-
CD34pos.CD117pos.frq	1.0106	(0.9923-1.0275)	0.2293	-	-	-
CD34neg.CD117pos.frq	0.9946	(0.9621-1.0177)	0.6943	-	-	-
CD34pos.CLEC12Apos.frq	1.0138	(0.9936-1.0316)	0.1448	-	-	-
CD34neg.CLEC12Apos.frq	0.9933	(0.9668-1.0131)	0.5626	-	-	-
leukocytter	0.9813	(0.9486-0.9996)	0.1446	-	-	-

Table 20: Poisson regression result for IDH2

	Risk Ratio	95% CI	P-Value	Risk Ratio	95% CI	P-Value
CD34pos.frq	1.0029	(0.988-1.0166)	0.6883	0.9915	(0.9743-1.0099)	0.3410
CD117pos.frq	1.0096	(0.9952-1.024)	0.1856	1.0176	(0.9926-1.0432)	0.1686
CD123pos.frq	0.9938	(0.9742-1.0105)	0.4999	0.9820	(0.9592-1.0027)	0.1034
CLEC12Apos.frq	1.0064	(0.9924-1.02)	0.3552	1.0055	(0.9858-1.0255)	0.5830
CD34pos.CD123pos.frq	1.0034	(0.9787-1.0238)	0.7637	-	-	-
CD34neg.CD123pos.frq	0.9906	(0.966-1.0091)	0.3827	-	-	-
CD34pos.CD117pos.frq	1.0064	(0.9914-1.0203)	0.3767	-	-	-
CD34neg.CD117pos.frq	1.0059	(0.9868-1.0216)	0.4990	-	-	-
CD34pos.CLEC12Apos.frq	1.0076	(0.9903-1.0227)	0.3499	-	-	-
CD34neg.CLEC12Apos.frq	1.0021	(0.9855-1.016)	0.7860	-	-	-
leukocytter	0.9962	(0.9841-1.0042)	0.4488	-	-	-

Immunophenotypes vs risk classification

```
riskFlowResults <- uniMultiRR(ys = risks,
                               xs = flows,
                               data = work.data,
                               nDigits = 2,
                               multixs = flows[1:4])
```

Table 21: Poisson regression result for MRC.Favorable

	Risk Ratio	95% CI	P-Value	Risk Ratio	95% CI	P-Value
CD34pos.frq	1.02	(1-1.03)	0.03	1.02	(1-1.04)	0.07
CD117pos.frq	1.01	(0.99-1.03)	0.23	0.99	(0.96-1.01)	0.28
CD123pos.frq	1.00	(0.98-1.02)	0.65	0.99	(0.97-1.02)	0.67
CLEC12Apos.frq	1.01	(1-1.03)	0.07	1.02	(1-1.04)	0.10
CD34pos.CD123pos.frq	1.02	(0.99-1.04)	0.13	-	-	-
CD34neg.CD123pos.frq	1.00	(0.98-1.02)	0.66	-	-	-
CD34pos.CD117pos.frq	1.02	(1-1.03)	0.05	-	-	-
CD34neg.CD117pos.frq	0.99	(0.95-1.01)	0.43	-	-	-
CD34pos.CLEC12Apos.frq	1.03	(1.01-1.04)	0.00	-	-	-
CD34neg.CLEC12Apos.frq	0.98	(0.95-1.01)	0.22	-	-	-
leukocytter	1.00	(0.99-1.01)	0.83	-	-	-

Table 22: Poisson regression result for MRC.Intermediate

	Risk Ratio	95% CI	P-Value	Risk Ratio	95% CI	P-Value
CD34pos.frq	1.00	(0.99-1)	0.24	1	(0.99-1)	0.19
CD117pos.frq	1.00	(0.99-1.01)	0.87	1	(0.99-1.01)	0.83
CD123pos.frq	1.00	(1-1.01)	0.25	1	(0.99-1.01)	0.66
CLEC12Apos.frq	1.00	(1-1.01)	0.27	1	(0.99-1.01)	0.53
CD34pos.CD123pos.frq	1.00	(0.99-1.01)	0.87	-	-	-
CD34neg.CD123pos.frq	1.00	(1-1.01)	0.32	-	-	-
CD34pos.CD117pos.frq	1.00	(0.99-1)	0.30	-	-	-
CD34neg.CD117pos.frq	1.00	(1-1.01)	0.15	-	-	-
CD34pos.CLEC12Apos.frq	1.00	(0.99-1)	0.42	-	-	-

	Risk Ratio	95% CI	P-Value	Risk Ratio	95% CI	P-Value
CD34neg.CLEC12Apos.frq	1.01	(1-1.01)	0.05	-	-	-
leukocyttter	1.00	(1-1)	0.37	-	-	-

Table 23: Poisson regression result for MRC.Adverse

	Risk Ratio	95% CI	P-Value	Risk Ratio	95% CI	P-Value
CD34pos.frq	1.00	(0.99-1.01)	0.41	1.01	(0.99-1.03)	0.18
CD117pos.frq	0.99	(0.98-1.01)	0.30	1.00	(0.98-1.02)	0.88
CD123pos.frq	0.98	(0.96-0.99)	0.02	0.99	(0.97-1.01)	0.40
CLEC12Apos.frq	0.98	(0.96-0.99)	0.00	0.98	(0.96-1)	0.02
CD34pos.CD123pos.frq	0.99	(0.97-1.01)	0.52	-	-	-
CD34neg.CD123pos.frq	0.98	(0.95-1)	0.04	-	-	-
CD34pos.CD117pos.frq	1.00	(0.99-1.01)	0.48	-	-	-
CD34neg.CD117pos.frq	0.98	(0.95-1)	0.04	-	-	-
CD34pos.CLEC12Apos.frq	0.99	(0.98-1.01)	0.46	-	-	-
CD34neg.CLEC12Apos.frq	0.97	(0.95-0.99)	0.01	-	-	-
leukocyttter	0.99	(0.98-1)	0.08	-	-	-

ROC analysis

Immunophenotypes vs mutational aberrations

```
ROCResults <- list()
for(y in mutations){
  tempList <- list()
  for(x in flows){
    test <- work.data[[x]]
    stat <- work.data[[y]]
    notNA <- (is.na(stat) == FALSE & is.na(test) == FALSE)

    tempRoc <- ROC(stat = stat[notNA],
                  test = test[notNA],
                  plot = NULL)
    if(tempRoc$AUC < 0.5){
      tempRoc <- ROC(stat = stat[notNA],
                    test = -test[notNA],
                    plot = NULL)
    }

    mx <- which.max(tempRoc$res$sens + tempRoc$res$spec)
    results <- c(tempRoc$AUC, as.numeric(tempRoc$res[mx,]))
    names(results) <- c("AUC", "Sens", "Spec", "PPV", "NPV", "cut")
    tempList[[x]] <- results[-c(4,5)] # Drop PPV and NPV - numbers are incorrect
  }
  dat <- data.frame(do.call(rbind, tempList))
  dat$mut_direction <- ifelse(dat$cut < 0, "below_cut", "above_cut")
  dat$cut <- abs(dat$cut)
  ROCResults[[y]] <- dat
}
```



```
rm(tempList, dat)
}
```

Table 24: ROC analysis for FLT3

	AUC	Sens	Spec	cut	mut_direction
CD34pos.frq	0.554	1.000	0.153	0.569	above_cut
CD117pos.frq	0.587	0.571	0.605	37.026	above_cut
CD123pos.frq	0.675	0.952	0.395	9.078	above_cut
CLEC12Apos.frq	0.701	0.714	0.609	30.868	above_cut
CD34pos.CD123pos.frq	0.610	0.429	0.791	19.406	above_cut
CD34neg.CD123pos.frq	0.699	0.810	0.577	4.349	above_cut
CD34pos.CD117pos.frq	0.553	0.952	0.200	0.494	above_cut
CD34neg.CD117pos.frq	0.583	0.952	0.302	0.418	above_cut
CD34pos.CLEC12Apos.frq	0.630	0.476	0.800	25.308	above_cut
CD34neg.CLEC12Apos.frq	0.569	0.238	0.935	68.492	above_cut
leukocytter	0.612	0.238	0.953	99.000	above_cut

Table 25: ROC analysis for NPM1

	AUC	Sens	Spec	cut	mut_direction
CD34pos.frq	0.863	0.821	0.822	2.078	below_cut
CD117pos.frq	0.657	0.538	0.756	12.485	below_cut
CD123pos.frq	0.550	0.410	0.716	26.113	above_cut
CLEC12Apos.frq	0.581	0.282	0.888	64.884	above_cut
CD34pos.CD123pos.frq	0.798	0.846	0.721	1.933	below_cut
CD34neg.CD123pos.frq	0.549	0.359	0.919	0.185	below_cut
CD34pos.CD117pos.frq	0.858	0.846	0.777	2.208	below_cut
CD34neg.CD117pos.frq	0.651	0.641	0.726	4.781	above_cut
CD34pos.CLEC12Apos.frq	0.803	0.769	0.777	1.515	below_cut
CD34neg.CLEC12Apos.frq	0.798	0.872	0.655	6.944	above_cut
leukocytter	0.653	0.579	0.735	21.100	above_cut

Table 26: ROC analysis for FLT3_NPM1

	AUC	Sens	Spec	cut	mut_direction
CD34pos.frq	0.708	0.895	0.571	13.832	below_cut
CD117pos.frq	0.564	1.000	0.171	2.567	above_cut
CD123pos.frq	0.767	0.789	0.724	24.575	above_cut
CLEC12Apos.frq	0.633	0.789	0.562	27.569	above_cut
CD34pos.CD123pos.frq	0.589	0.895	0.355	12.904	below_cut
CD34neg.CD123pos.frq	0.726	0.789	0.641	5.905	above_cut
CD34pos.CD117pos.frq	0.683	0.895	0.567	11.915	below_cut
CD34neg.CD117pos.frq	0.847	0.842	0.783	9.263	above_cut
CD34pos.CLEC12Apos.frq	0.708	0.842	0.581	5.640	below_cut
CD34neg.CLEC12Apos.frq	0.820	0.789	0.783	20.003	above_cut
leukocytter	0.769	0.895	0.600	10.100	above_cut

Table 27: ROC analysis for WT1

	AUC	Sens	Spec	cut	mut_direction
CD34pos.frq	0.568	0.25	0.983	0.133	below_cut
CD117pos.frq	0.504	0.25	0.906	1.919	below_cut
CD123pos.frq	0.526	0.50	0.867	48.617	above_cut
CLEC12Apos.frq	0.522	0.25	0.989	0.652	below_cut
CD34pos.CD123pos.frq	0.506	0.25	0.972	52.841	above_cut
CD34neg.CD123pos.frq	0.517	1.00	0.249	16.867	below_cut
CD34pos.CD117pos.frq	0.526	0.75	0.431	20.936	below_cut
CD34neg.CD117pos.frq	0.630	1.00	0.398	0.892	above_cut
CD34pos.CLEC12Apos.frq	0.511	0.25	0.972	0.105	below_cut
CD34neg.CLEC12Apos.frq	0.530	1.00	0.227	0.740	above_cut
leukocytter	0.591	1.00	0.296	2.670	above_cut

Table 28: ROC analysis for cKIT

	AUC	Sens	Spec	cut	mut_direction
CD34pos.frq	0.639	1.000	0.543	21.212	above_cut
CD117pos.frq	0.540	1.000	0.413	23.461	above_cut
CD123pos.frq	0.627	1.000	0.500	12.485	above_cut
CLEC12Apos.frq	0.674	1.000	0.576	30.346	above_cut
CD34pos.CD123pos.frq	0.732	1.000	0.701	15.074	above_cut
CD34neg.CD123pos.frq	0.783	1.000	0.723	12.762	above_cut
CD34pos.CD117pos.frq	0.627	1.000	0.565	19.423	above_cut
CD34neg.CD117pos.frq	0.587	0.667	0.696	0.607	below_cut
CD34pos.CLEC12Apos.frq	0.788	1.000	0.712	20.793	above_cut
CD34neg.CLEC12Apos.frq	0.598	0.667	0.750	0.889	below_cut
leukocytter	0.702	1.000	0.610	12.100	above_cut

Table 29: ROC analysis for CEBPA

	AUC	Sens	Spec	cut	mut_direction
CD34pos.frq	0.748	0.636	0.803	47.115	above_cut
CD117pos.frq	0.708	0.545	0.850	61.514	above_cut
CD123pos.frq	0.544	0.818	0.422	10.932	above_cut
CLEC12Apos.frq	0.799	0.909	0.775	47.727	above_cut
CD34pos.CD123pos.frq	0.731	0.636	0.838	21.942	above_cut
CD34neg.CD123pos.frq	0.652	0.818	0.595	5.095	above_cut
CD34pos.CD117pos.frq	0.776	0.545	0.913	59.162	above_cut
CD34neg.CD117pos.frq	0.519	1.000	0.202	21.815	below_cut
CD34pos.CLEC12Apos.frq	0.834	0.818	0.769	22.493	above_cut
CD34neg.CLEC12Apos.frq	0.507	0.273	0.925	0.317	below_cut
leukocytter	0.569	1.000	0.287	2.500	above_cut

Table 30: ROC analysis for IDH1

	AUC	Sens	Spec	cut	mut_direction
CD34pos.frq	0.588	0.706	0.542	18.947	above_cut
CD117pos.frq	0.564	0.471	0.708	47.280	above_cut
CD123pos.frq	0.603	0.471	0.881	45.482	above_cut
CLEC12Apos.frq	0.536	0.529	0.631	37.932	above_cut
CD34pos.CD123pos.frq	0.661	0.412	0.940	35.080	above_cut
CD34neg.CD123pos.frq	0.539	0.235	0.917	41.520	above_cut
CD34pos.CD117pos.frq	0.593	0.588	0.661	23.593	above_cut
CD34neg.CD117pos.frq	0.501	0.941	0.190	25.189	below_cut
CD34pos.CLEC12Apos.frq	0.619	0.706	0.619	11.434	above_cut
CD34neg.CLEC12Apos.frq	0.549	0.647	0.536	4.422	below_cut
leukocyttter	0.679	0.824	0.530	7.450	below_cut

Table 31: ROC analysis for IDH2

	AUC	Sens	Spec	cut	mut_direction
CD34pos.frq	0.517	0.296	0.872	0.747	below_cut
CD117pos.frq	0.595	0.889	0.391	18.168	above_cut
CD123pos.frq	0.523	0.815	0.276	36.076	below_cut
CLEC12Apos.frq	0.581	0.963	0.224	7.690	above_cut
CD34pos.CD123pos.frq	0.515	0.148	0.962	0.118	below_cut
CD34neg.CD123pos.frq	0.601	0.630	0.654	1.987	below_cut
CD34pos.CD117pos.frq	0.529	0.185	0.936	65.303	above_cut
CD34neg.CD117pos.frq	0.584	0.963	0.263	0.336	above_cut
CD34pos.CLEC12Apos.frq	0.527	0.222	0.897	46.758	above_cut
CD34neg.CLEC12Apos.frq	0.515	0.407	0.737	18.341	above_cut
leukocyttter	0.585	0.731	0.490	9.600	below_cut

Immunophenotypes vs risk classification

```

ROCResults <- list()
for(y in risks){
  tempList <- list()
  for(x in flows){
    test <- work.data[[x]]
    stat <- work.data[[y]]
    notNA <- (is.na(stat) == FALSE & is.na(test) == FALSE)

    tempRoc <- ROC(stat = stat[notNA],
                  test = test[notNA],
                  plot = NULL)
    if(tempRoc$AUC < 0.5){
      tempRoc <- ROC(stat = stat[notNA],
                    test = -test[notNA],
                    plot = NULL)
    }

    mx <- which.max(tempRoc$res$sens + tempRoc$res$spec)
    results <- c(tempRoc$AUC, as.numeric(tempRoc$res[mx,]))
  }
}

```

```

names(results) <- c("AUC", "Sens", "Spec", "PPV", "NPV", "cut")
tempList[[x]] <- results[-c(4,5)] # Drop PPV and NPV - numbers are incorrect
}
dat <- data.frame(do.call(rbind, tempList))
dat$mut_direction <- ifelse(dat$cut < 0, "below_cut", "above_cut")
dat$cut <- abs(dat$cut)
ROCResults[[y]] <- dat

rm(tempList, dat)
}

```

Table 32: ROC analysis for MRC.Favorable

	AUC	Sens	Spec	cut	mut_direction
CD34pos.frq	0.677	0.75	0.656	25.349	above_cut
CD117pos.frq	0.602	0.85	0.456	23.461	above_cut
CD123pos.frq	0.567	0.80	0.465	10.966	above_cut
CLEC12Apos.frq	0.648	0.85	0.531	23.472	above_cut
CD34pos.CD123pos.frq	0.694	0.90	0.552	4.689	above_cut
CD34neg.CD123pos.frq	0.567	0.55	0.697	9.863	above_cut
CD34pos.CD117pos.frq	0.677	0.80	0.618	19.423	above_cut
CD34neg.CD117pos.frq	0.578	0.80	0.515	1.920	below_cut
CD34pos.CLEC12Apos.frq	0.761	0.80	0.763	20.793	above_cut
CD34neg.CLEC12Apos.frq	0.646	0.70	0.639	2.468	below_cut
leukocytter	0.538	0.75	0.485	5.780	above_cut

Table 33: ROC analysis for MRC.Intermediate

	AUC	Sens	Spec	cut	mut_direction
CD34pos.frq	0.632	0.642	0.635	21.622	below_cut
CD117pos.frq	0.505	0.316	0.797	46.699	above_cut
CD123pos.frq	0.566	0.364	0.824	24.575	above_cut
CLEC12Apos.frq	0.576	0.412	0.757	34.908	above_cut
CD34pos.CD123pos.frq	0.609	0.529	0.743	2.988	below_cut
CD34neg.CD123pos.frq	0.542	0.631	0.514	2.012	above_cut
CD34pos.CD117pos.frq	0.620	0.513	0.784	6.139	below_cut
CD34neg.CD117pos.frq	0.619	0.374	0.905	8.452	above_cut
CD34pos.CLEC12Apos.frq	0.613	0.455	0.770	2.832	below_cut
CD34neg.CLEC12Apos.frq	0.654	0.439	0.865	11.249	above_cut
leukocytter	0.569	0.384	0.797	17.300	above_cut

Table 34: ROC analysis for MRC.Adverse

	AUC	Sens	Spec	cut	mut_direction
CD34pos.frq	0.587	0.796	0.435	5.691	above_cut
CD117pos.frq	0.550	0.852	0.319	47.280	below_cut
CD123pos.frq	0.611	0.852	0.357	24.575	below_cut
CLEC12Apos.frq	0.658	0.852	0.473	30.346	below_cut
CD34pos.CD123pos.frq	0.551	0.852	0.348	1.062	above_cut
CD34neg.CD123pos.frq	0.581	0.852	0.333	11.673	below_cut

	AUC	Sens	Spec	cut	mut_direction
CD34pos.CD117pos.frq	0.573	0.759	0.478	6.086	above_cut
CD34neg.CD117pos.frq	0.613	0.926	0.353	8.975	below_cut
CD34pos.CLEC12Apos.frq	0.527	0.833	0.333	0.957	above_cut
CD34neg.CLEC12Apos.frq	0.628	0.889	0.415	12.371	below_cut
leukocyttter	0.601	0.722	0.488	9.810	below_cut

Survival analysis

Remove leukocyttter from covariates for survival analysis

```
flows <- flows[-grep("leukocyttter",flows)]
```

These analyses are performed separately for patients with / without curative intended treatment

Curative intent	N
0	122
1	144

Curative Intent

Perform Uni and multivariate cox regression with continous markers

```
work.data$MRC <- factor(work.data$`MRC classification`)
coxData <- work.data[,c("Alder", "MRC", "FLT3", "NPM1", "FLT3_NPM1", flows)]
coxData$OS <- Surv(work.data$OS, work.data$`Kaplan-Meier overlevelse`)
coxDataC <- coxData[work.data$`Kurativt sigte?` == 1,]
formula = as.formula(paste("OS~", paste(flows, collapse = "+")))
coxResults <- coxR(formula, data = coxDataC, multixs = c(flows[1:4]),
  fix = c("Alder", "MRC", "FLT3", "NPM1", "FLT3_NPM1"))
```

Table 36: Results of uni (left) and multivariate (right) Cox regressions adjusted for age, MRC, FLT3, NMP1, and FLT3_NPM1. n = 141, number of events = 81

	Hazard ratio	95% CI	P-Value	Hazard ratio	95% CI	P-Value
CD34pos.frq	1	(0.99-1.01)	0.39	1.01	(1-1.03)	0.059
CD117pos.frq	1	(0.99-1.01)	0.6	0.99	(0.98-1)	0.12
CD123pos.frq	1.01	(1-1.02)	0.061	1.03	(1.01-1.04)	0.00052
CLEC12Apos.frq	0.99	(0.98-1)	0.18	0.98	(0.97-1)	0.012
CD34pos.CD123pos.frq	1.01	(1-1.03)	0.15	-	-	-
CD34neg.CD123pos.frq	1	(0.99-1.01)	0.62	-	-	-
CD34pos.CD117pos.frq	1	(0.99-1.01)	0.84	-	-	-
CD34neg.CD117pos.frq	0.99	(0.98-1.01)	0.38	-	-	-
CD34pos.CLEC12Apos.frq	1	(0.99-1.01)	0.67	-	-	-
CD34neg.CLEC12Apos.frq	0.99	(0.98-1.01)	0.33	-	-	-

Perform cox regressions with dichotomized values, i.e. expressions above or below median.

```
## Create data
coxDataDic <- data.frame(lapply(subset(work.data, select = c(flows)), splitMedian))
```

```

coxDataDic$OS <- Surv(work.data$OS, work.data$`Kaplan-Meier overlevelse`)
coxDataDic <- cbind(coxDataDic, work.data[,c("Alder", "MRC", "FLT3", "NPM1", "FLT3_NPM1")])

## Subset to curative intent
coxDataDicC <- coxDataDic[work.data$`Kurativt sigte?` == 1,]
formula = as.formula(paste("OS~", paste(flows, collapse = "+")))
coxResultsDic <- coxR(formula, data = coxDataDicC, multixs = c(flows[1:4]),
  fix = c("Alder", "MRC", "FLT3", "NPM1", "FLT3_NPM1"))

```

Table 37: Results of uni (left) and multivariate (right) Cox regressions adjusted for age, MRC, FLT3, NPM1, and FLT3_NPM1 with dichotomized expressions. n = 141, number of events = 81

	Hazard ratio	95% CI	P-Value	Hazard ratio	95% CI	P-Value
CD34pos.frqLow	1	-	-	-	-	-
CD34pos.frqHigh	0.94	(0.57-1.56)	0.82	1.06	(0.59-1.9)	0.85
CD117pos.frqLow	1	-	-	-	-	-
CD117pos.frqHigh	0.88	(0.56-1.39)	0.59	0.99	(0.57-1.72)	0.97
CD123pos.frqLow	1	-	-	-	-	-
CD123pos.frqHigh	0.89	(0.55-1.43)	0.62	1.05	(0.59-1.85)	0.87
CLEC12Apos.frqLow	1	-	-	-	-	-
CLEC12Apos.frqHigh	0.73	(0.44-1.19)	0.21	0.7	(0.37-1.31)	0.27
CD34pos.CD123pos.frqLow	1	-	-	-	-	-
CD34pos.CD123pos.frqHigh	1.19	(0.72-1.97)	0.5	-	-	-
CD34neg.CD123pos.frqLow	1	-	-	-	-	-
CD34neg.CD123pos.frqHigh	0.89	(0.55-1.43)	0.62	-	-	-
CD34pos.CD117pos.frqLow	1	-	-	-	-	-
CD34pos.CD117pos.frqHigh	1.1	(0.66-1.81)	0.72	-	-	-
CD34neg.CD117pos.frqLow	1	-	-	-	-	-
CD34neg.CD117pos.frqHigh	0.79	(0.48-1.31)	0.37	-	-	-
CD34pos.CLEC12Apos.frqLow	1	-	-	-	-	-
CD34pos.CLEC12Apos.frqHigh	1	(0.62-1.61)	0.99	-	-	-
CD34neg.CLEC12Apos.frqLow	1	-	-	-	-	-
CD34neg.CLEC12Apos.frqHigh	0.73	(0.44-1.24)	0.25	-	-	-

Not Curative Intent

Perform Uni and multivariate cox regression with continous markers.

```

coxDataNC <- coxData[work.data$`Kurativt sigte?` == 0,]
formula = as.formula(paste("OS~", paste(flows, collapse = "+")))
coxResults <- coxR(formula, data = coxDataNC, multixs = c(flows[1:4]),
  fix = c("Alder", "MRC", "FLT3", "NPM1", "FLT3_NPM1"))

```

Table 38: Results of uni (left) and multivariate (right) Cox regressions adjusted for age, MRC, FLT3, NPM1, and FLT3_NPM1. n = 87, number of events = 83

	Hazard ratio	95% CI	P-Value	Hazard ratio	95% CI	P-Value
CD34pos.frq	1	(0.99-1.01)	0.86	0.99	(0.98-1.01)	0.41
CD117pos.frq	1	(0.99-1.01)	0.85	1	(0.98-1.02)	0.97
CD123pos.frq	1	(0.99-1.01)	0.82	0.99	(0.98-1.01)	0.24
CLEC12Apos.frq	1.01	(1-1.02)	0.11	1.02	(1-1.03)	0.032

	Hazard ratio	95% CI	P-Value	Hazard ratio	95% CI	P-Value
CD34pos.CD123pos.frq	1	(0.98-1.02)	0.97	-	-	-
CD34neg.CD123pos.frq	1	(0.99-1.02)	0.88	-	-	-
CD34pos.CD117pos.frq	1	(0.99-1.01)	0.84	-	-	-
CD34neg.CD117pos.frq	1	(0.98-1.02)	0.92	-	-	-
CD34pos.CLEC12Apos.frq	1.01	(1-1.03)	0.17	-	-	-
CD34neg.CLEC12Apos.frq	1	(0.99-1.02)	0.65	-	-	-

Perform cox regressions with dichotomized values, i.e. expressions above or below median.

```
coxDataDicNC <- coxDataDic[work.data$`Kurativt sigte?` == 0,]
formula = as.formula(paste("OS~", paste(flows, collapse = "+")))
coxResultsDic <- coxR(formula, data = coxDataDicNC, multixs = c(flows[1:4]),
  fix = c("Alder", "MRC", "FLT3", "NPM1", "FLT3_NPM1"))
```

Table 39: Results of uni (left) and multivariate (right) Cox regressions adjusted for age, MRC, FLT3, NPM1, and FLT3_NPM1, with dichotomized expressions. n = 87, number of events = 83

	Hazard ratio	95% CI	P-Value	Hazard ratio	95% CI	P-Value
CD34pos.frqLow	1	-	-	-	-	-
CD34pos.frqHigh	0.57	(0.32-1.01)	0.056	0.58	(0.31-1.09)	0.091
CD117pos.frqLow	1	-	-	-	-	-
CD117pos.frqHigh	0.89	(0.54-1.47)	0.66	0.93	(0.49-1.74)	0.81
CD123pos.frqLow	1	-	-	-	-	-
CD123pos.frqHigh	0.96	(0.6-1.55)	0.88	0.89	(0.52-1.53)	0.68
CLEC12Apos.frqLow	1	-	-	-	-	-
CLEC12Apos.frqHigh	1.24	(0.73-2.1)	0.43	1.38	(0.72-2.67)	0.33
CD34pos.CD123pos.frqLow	1	-	-	-	-	-
CD34pos.CD123pos.frqHigh	0.76	(0.47-1.23)	0.27	-	-	-
CD34neg.CD123pos.frqLow	1	-	-	-	-	-
CD34neg.CD123pos.frqHigh	0.95	(0.6-1.5)	0.82	-	-	-
CD34pos.CD117pos.frqLow	1	-	-	-	-	-
CD34pos.CD117pos.frqHigh	0.72	(0.41-1.26)	0.25	-	-	-
CD34neg.CD117pos.frqLow	1	-	-	-	-	-
CD34neg.CD117pos.frqHigh	0.86	(0.52-1.45)	0.58	-	-	-
CD34pos.CLEC12Apos.frqLow	1	-	-	-	-	-
CD34pos.CLEC12Apos.frqHigh	0.85	(0.5-1.42)	0.53	-	-	-
CD34neg.CLEC12Apos.frqLow	1	-	-	-	-	-
CD34neg.CLEC12Apos.frqHigh	0.73	(0.42-1.28)	0.27	-	-	-

Analysis of MFI

Pick data for analysis

```
mfi.col <- grep("MFI", names(work.data))
coxData <- work.data[, mfi.col] / 1000
coxData <- cbind(coxData, work.data[,c("Alder", "MRC", "FLT3", "NPM1", "FLT3_NPM1")])
coxData$OS <- Surv(work.data$OS, work.data$`Kaplan-Meier overlevelse`)
names(coxData) <- gsub("\\+ \\| ", "_", names(coxData))
```

Curative intent

```
coxDataC <- coxData[work.data$`Kurativt sigte?` == 1,]  
formula = as.formula(paste("OS~", paste(names(coxDataC)[1:4], collapse = "+")))  
coxResults <- coxR(formula, data = coxDataC, multixs = names(coxDataC)[1:4],  
  fix = c("Alder", "MRC", "FLT3", "NPM1", "FLT3_NPM1"))
```

Table 40: Results of uni (left) and multivariate (right) Cox regressions adjusted for age. n = 140, number of events = 80

	Hazard ratio	95% CI	P-Value	Hazard ratio	95% CI	P-Value
CD34_MFI	1.01	(0.99-1.02)	0.24	1.01	(1-1.03)	0.14
CD117_MFI	0.99	(0.98-1)	0.12	0.99	(0.98-1)	0.095
CD123_MFI	1.04	(0.96-1.12)	0.35	1.06	(0.98-1.14)	0.15
CLEC12A_MFI	0.96	(0.92-1.01)	0.11	0.96	(0.91-1.01)	0.12

Not curative intent

```
coxDataNC <- coxData[work.data$`Kurativt sigte?` == 0,]  
formula = as.formula(paste("OS~", paste(names(coxDataNC)[1:4], collapse = "+")))  
coxResults <- coxR(formula, data = coxDataNC, multixs = names(coxDataNC)[1:4],  
  fix = c("Alder", "MRC", "FLT3", "NPM1", "FLT3_NPM1"))
```

Table 41: Results of uni (left) and multivariate (right) Cox regressions adjusted for age. n = 86, number of events = 82

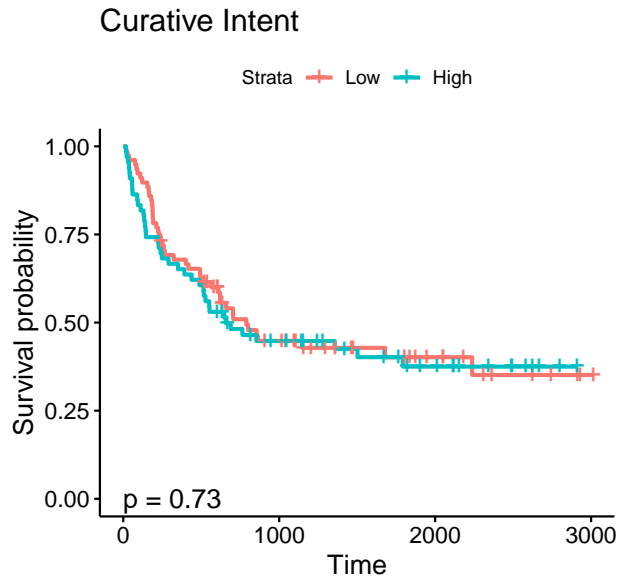
	Hazard ratio	95% CI	P-Value	Hazard ratio	95% CI	P-Value
CD34_MFI	1	(0.99-1.01)	0.9	1	(0.99-1.01)	0.94
CD117_MFI	1	(0.99-1.01)	0.89	1	(0.99-1.01)	0.87
CD123_MFI	0.97	(0.86-1.1)	0.65	1.01	(0.89-1.15)	0.82
CLEC12A_MFI	0.95	(0.89-1)	0.048	0.94	(0.89-1)	0.05

Kaplan-Meier plots

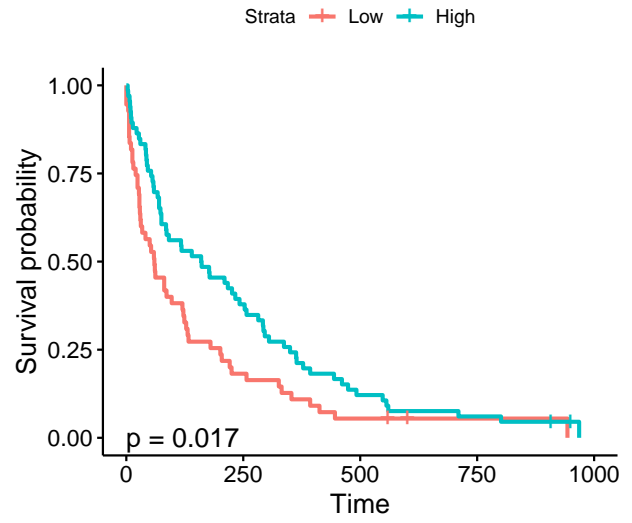
Split by median

```
## Warning: `select_()` is deprecated as of dplyr 0.7.0.  
## Please use `select()` instead.  
## This warning is displayed once every 8 hours.  
## Call `lifecycle::last_warnings()` to see where this warning was generated.
```

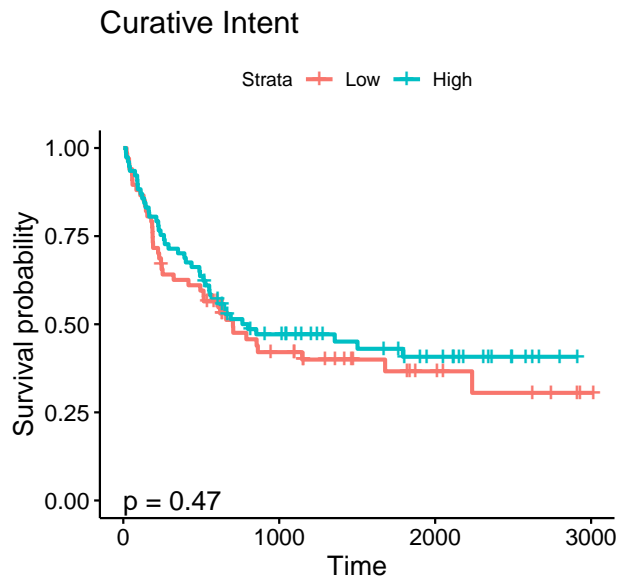

CD34pos vs Survival



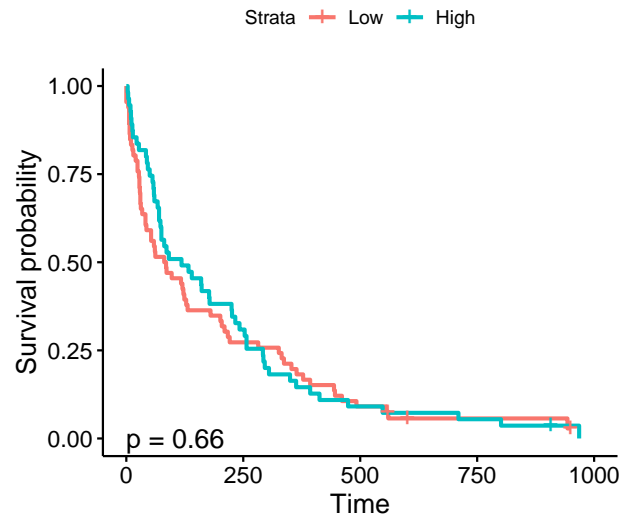
Not Curative Intent



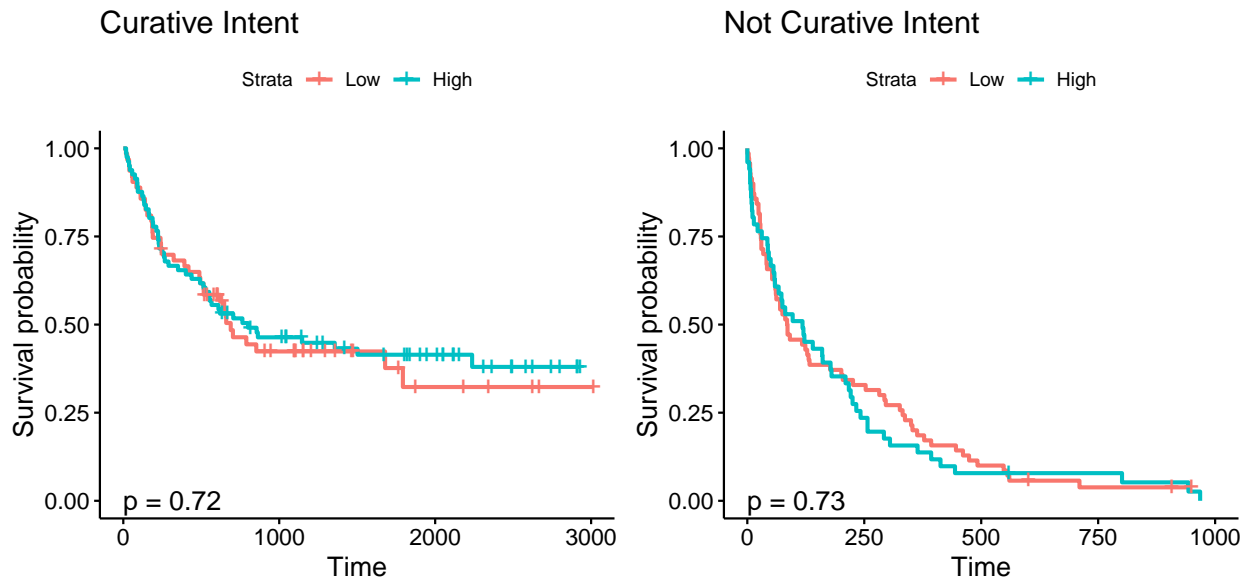
CD117pos vs Survival



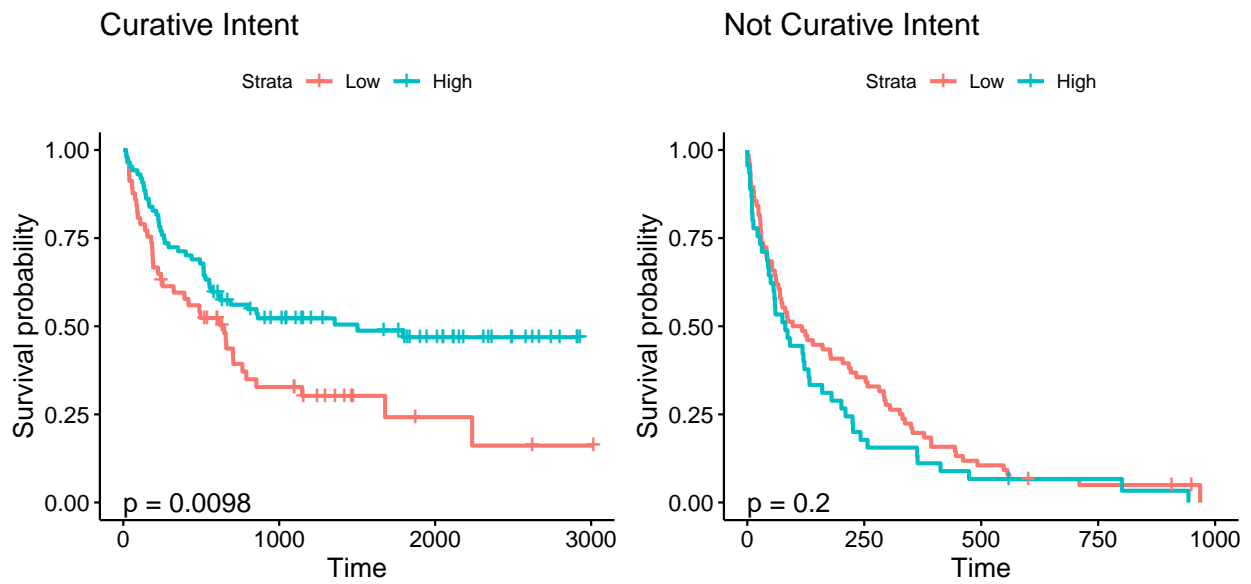
Not Curative Intent



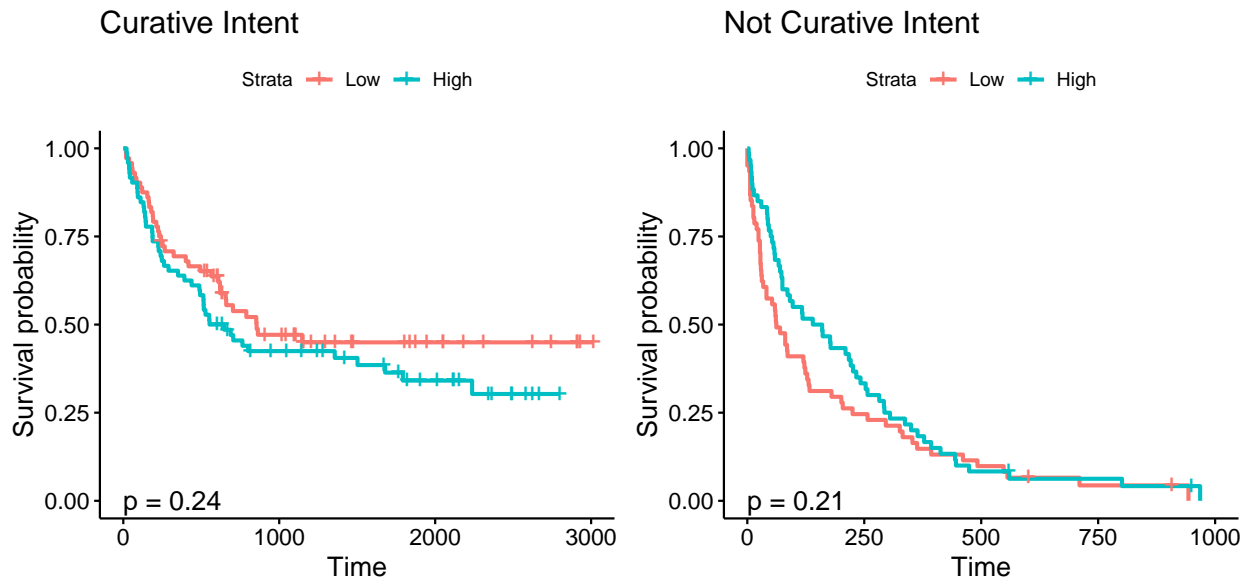
CD123pos vs Survival



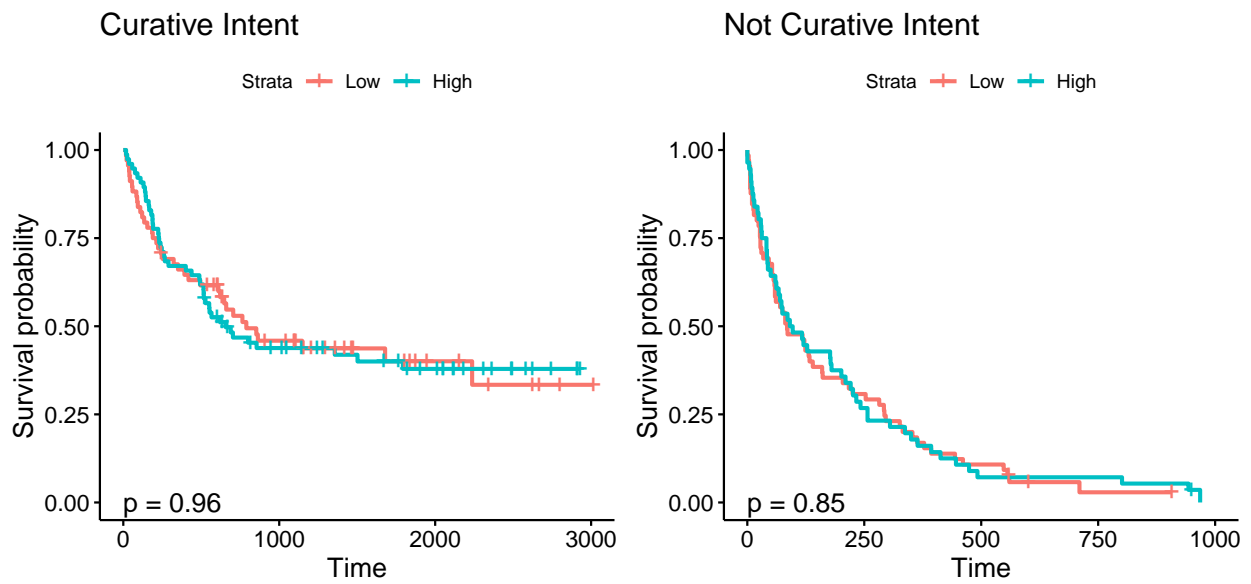
CLEC12Apos vs Survival



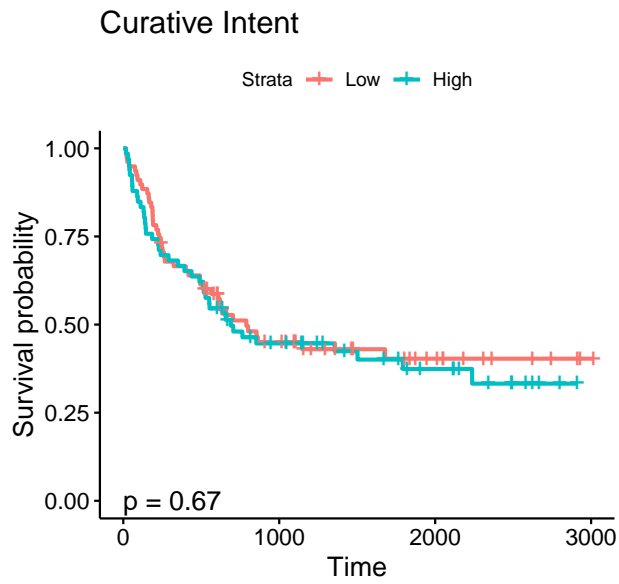
CD34pos.CD123pos vs Survival



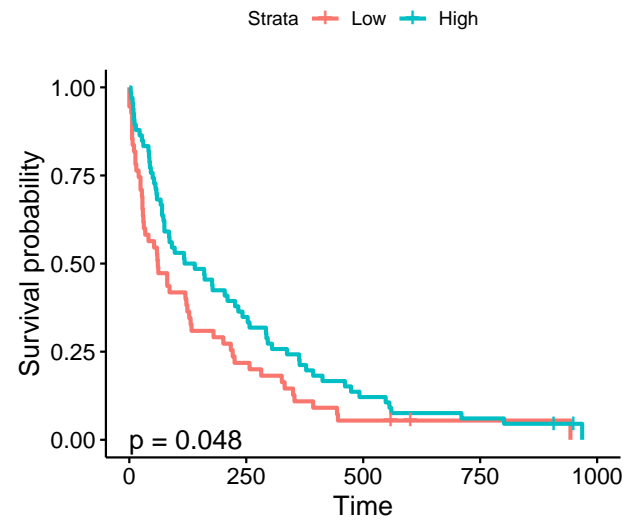
CD34neg.CD123pos vs Survival



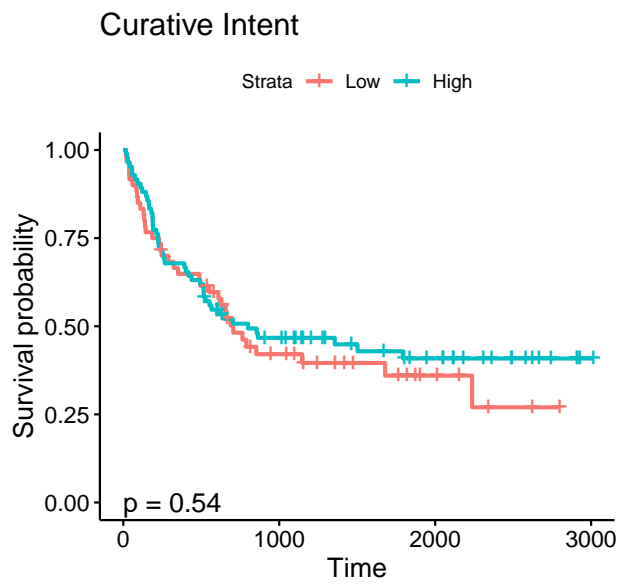
CD34pos.CD117pos vs Survival



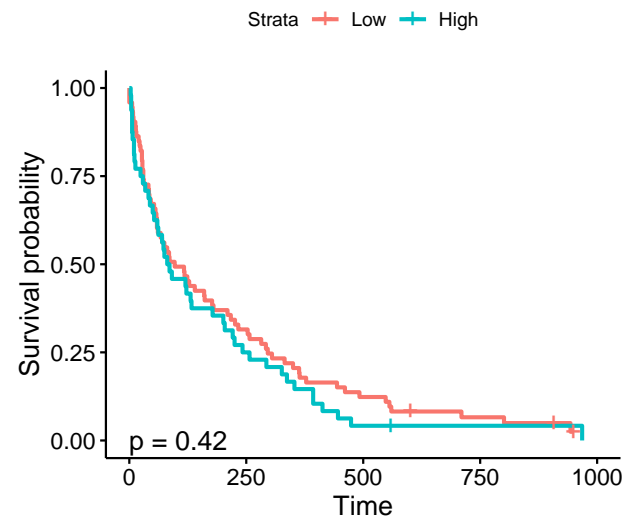
Not Curative Intent



CD34neg.CD117pos vs Survival



Not Curative Intent



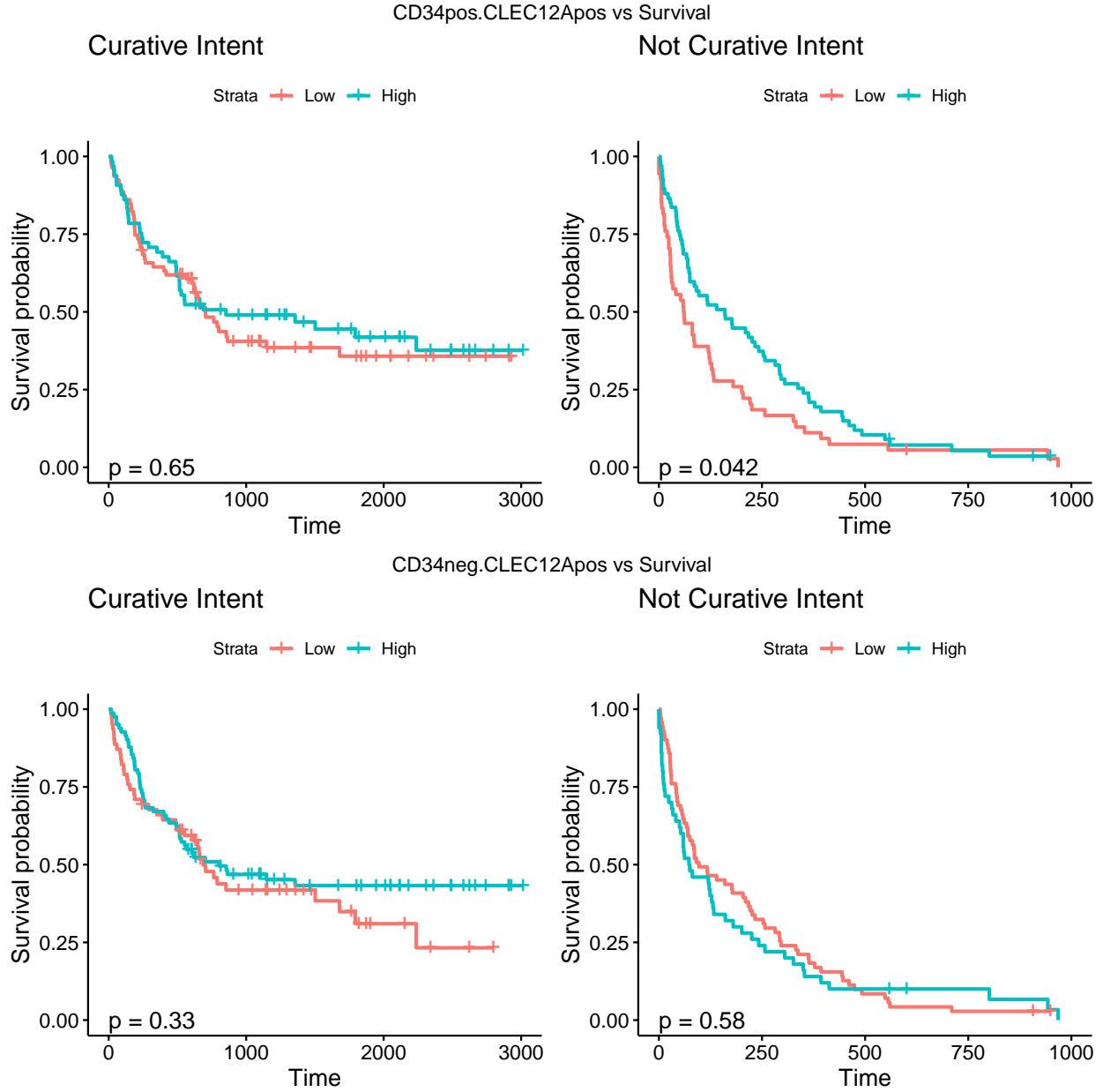
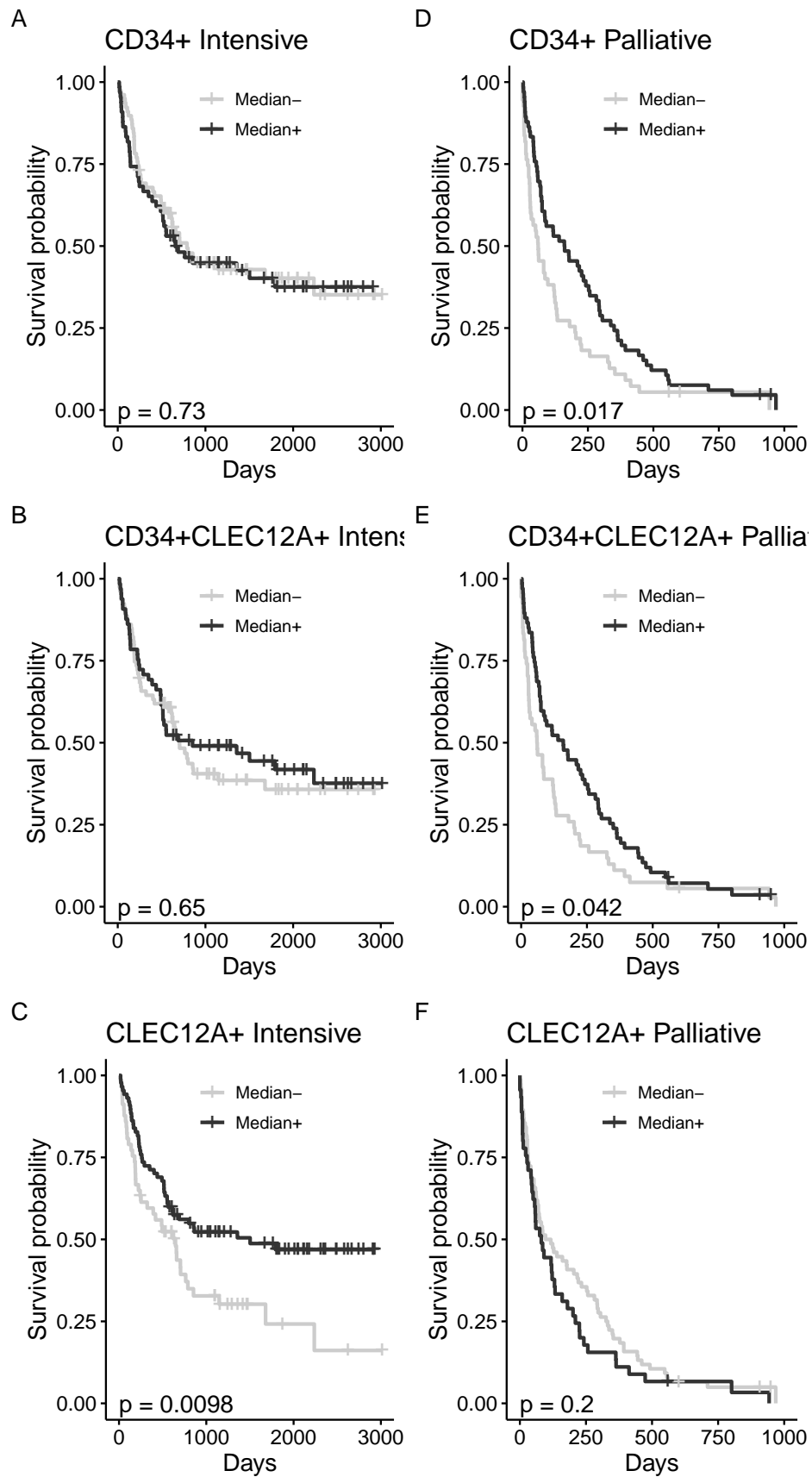


Table 42: Median survival times for flowmarker pos vs neg for patients with respectively intensive or palliative treatment intent

	Intensive median-	Intensive median+	Palliative median-	Palliative median+
CD34pos.frq	789	657	60.0	160.5
CD117pos.frq	703	800	83.0	118.0
CD123pos.frq	689	800	85.5	118.0
CLEC12Apos.frq	647	1502	107.0	81.0
CD34pos.CD123pos.frq	855	600	62.0	150.0
CD34neg.CD123pos.frq	789	657	85.0	94.0
CD34pos.CD117pos.frq	789	689	61.0	129.0
CD34neg.CD117pos.frq	705	800	97.0	83.5
CD34pos.CLEC12Apos.frq	703	853	60.5	160.0

	Intensive median-	Intensive median+	Palliative median-	Palliative median+
CD34neg.CLEC12Apos.frq	689	800	97.0	74.0



Split by quartiles

Generate data for KM plots

```
coxDataDic <- data.frame(lapply(subset(work.data, select = c(flows)), splitQuart))
coxDataDic$OS <- Surv(work.data$OS, work.data$`Kaplan-Meier overlevelse`)
coxDataDic$Alder <- work.data$Alder

coxDataDicC <- coxDataDic[work.data$`Kurativt sigte?` == 1,]
coxDataDicNC <- coxDataDic[work.data$`Kurativt sigte?` == 0,]

p.table.quart <- data.frame(quart.curative = rep(NA, length(flows)),
                           quart.noncurative = rep(NA, length(flows)))
row.names(p.table.quart) <- flows
for(flow in flows){
  formula <- as.formula(paste0("OS~",flow))
  curative <- surv_fit(formula, data = coxDataDicC)
  noncurative <- surv_fit(formula, data = coxDataDicNC)

  p.table.quart[flow, "quart.curative"] <- as.numeric(sub("p = ", "", surv_pvalue(curative)$pval))
  p.table.quart[flow, "quart.noncurative"] <- as.numeric(sub("p = ", "", surv_pvalue(noncurative)$pval))

  p1 <- ggsurvplot(curative,
                   risk.table = F,
                   data = coxDataDicC,
                   title = "Curative Intent",
                   pval = TRUE,
                   pval.coord = c(0, 0),
                   legend.labs = c("Q1", "Q2", "Q3", "Q4"))

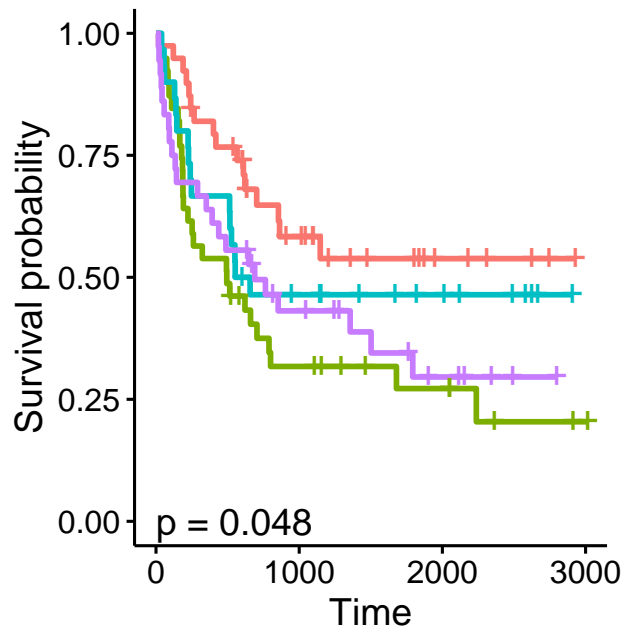
  p2 <- ggsurvplot(noncurative,
                   risk.table = F,
                   data = coxDataDicNC,
                   title = "Not Curative Intent",
                   pval = TRUE,
                   pval.coord = c(0, 0),
                   legend.labs = c("Q1", "Q2", "Q3", "Q4"))

  arrange_ggsurvplots(list(p1,p2), title = paste(sub(".frq","", flow), "vs Survival"))
}
```


CD34pos vs Survival

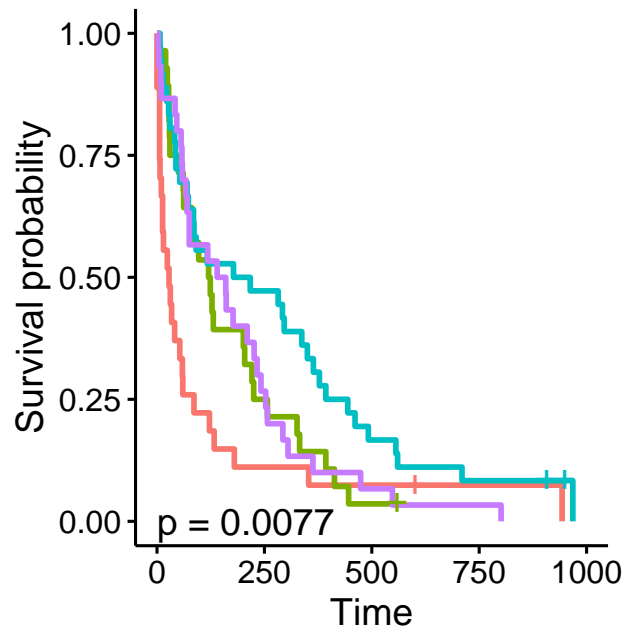
Curative Intent

Strata + Q1 + Q2 + Q3 + Q4



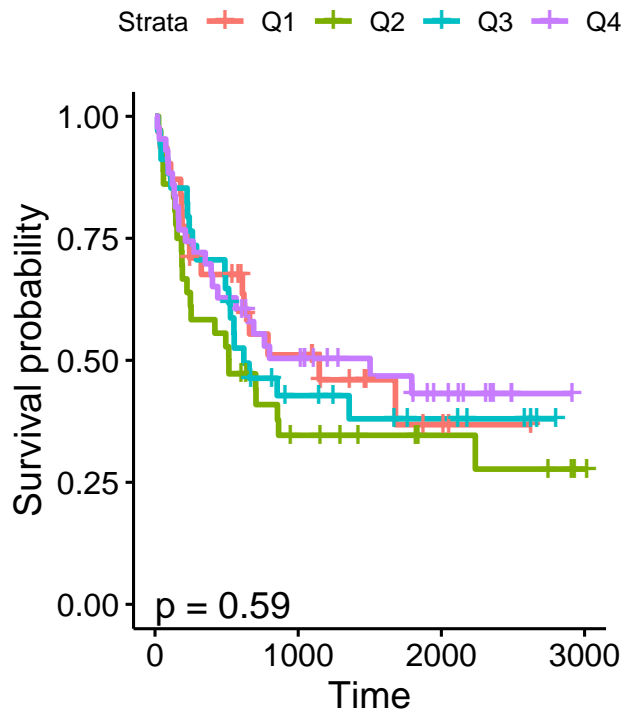
Not Curative Intent

Strata + Q1 + Q2 + Q3 + Q4

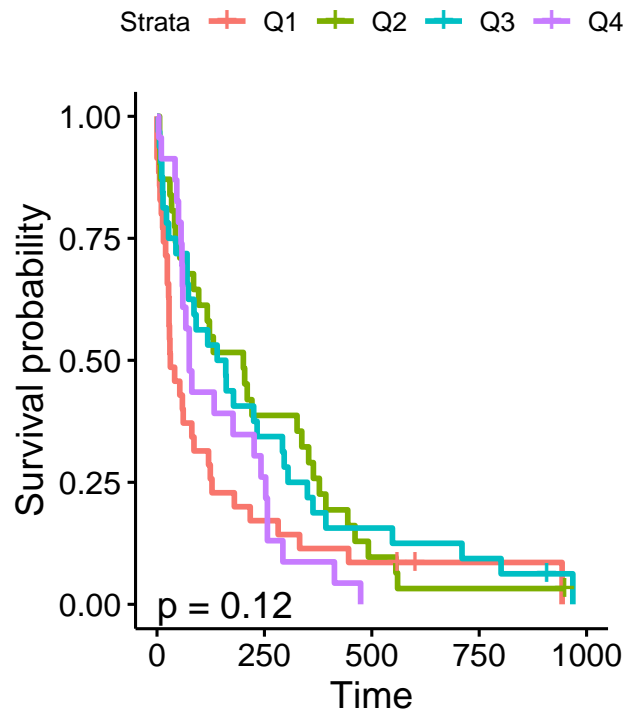


CD117pos vs Survival

Curative Intent



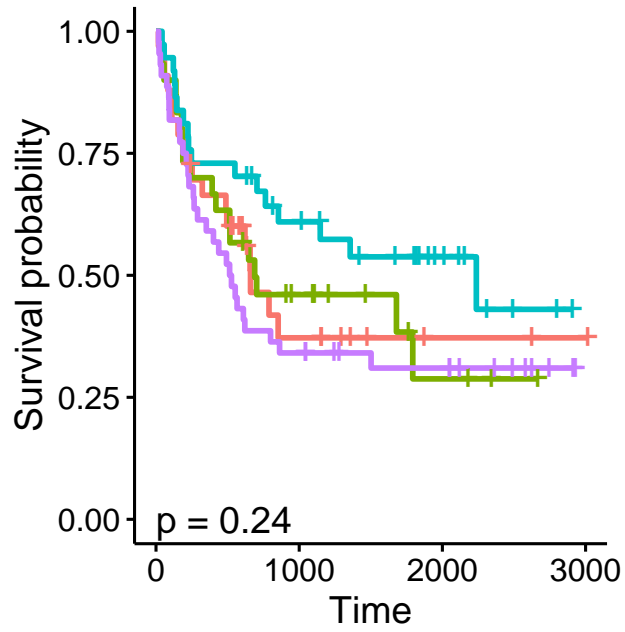
Not Curative Intent



CD123pos vs Survival

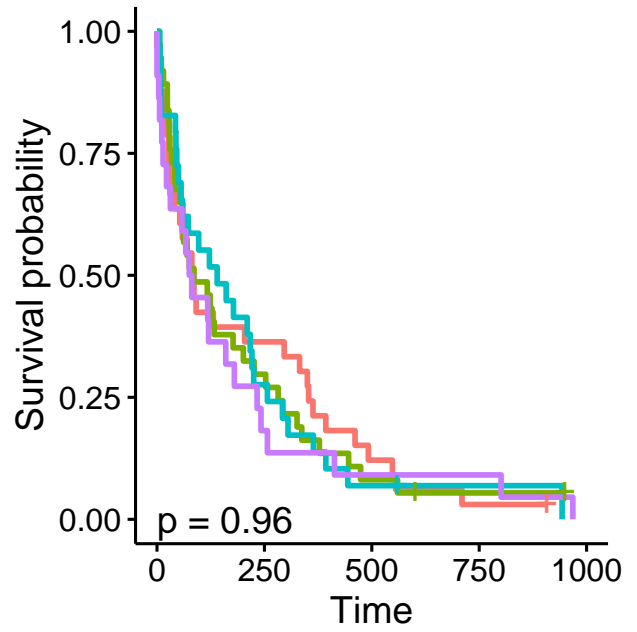
Curative Intent

Strata + Q1 + Q2 + Q3 + Q4



Not Curative Intent

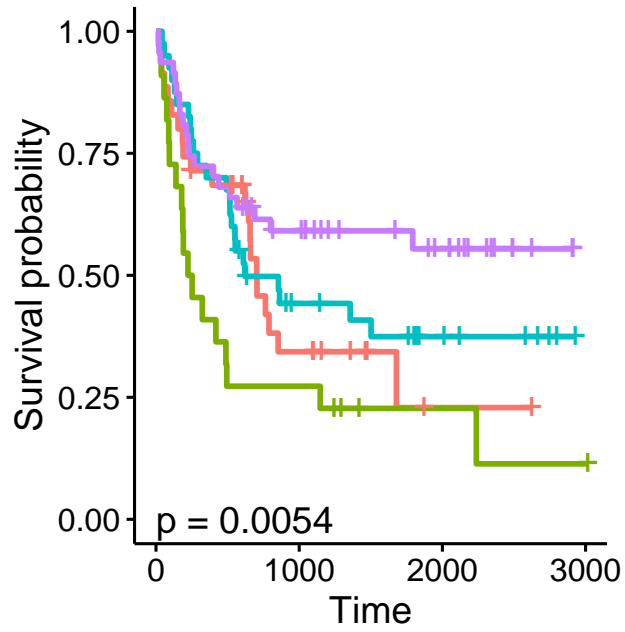
Strata + Q1 + Q2 + Q3 + Q4



CLEC12Apos vs Survival

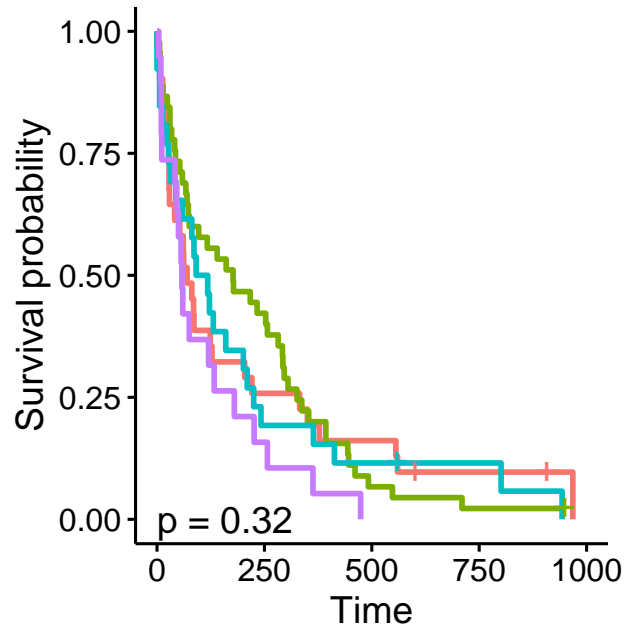
Curative Intent

Strata + Q1 + Q2 + Q3 + Q4



Not Curative Intent

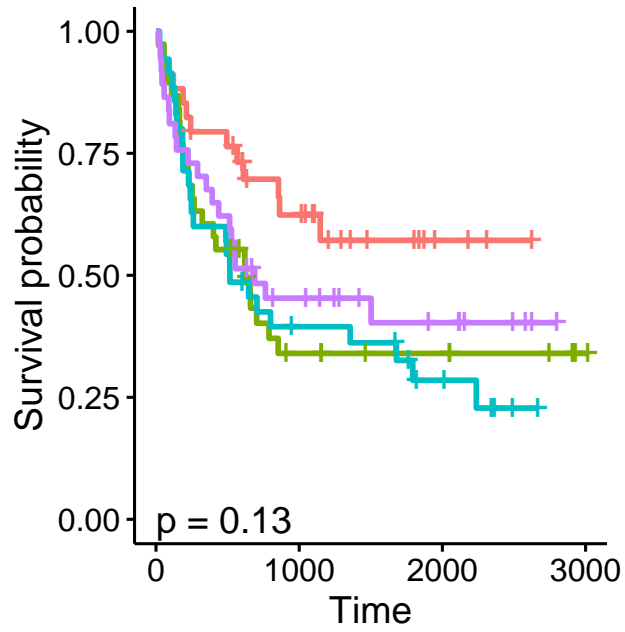
Strata + Q1 + Q2 + Q3 + Q4



CD34pos.CD123pos vs Survival

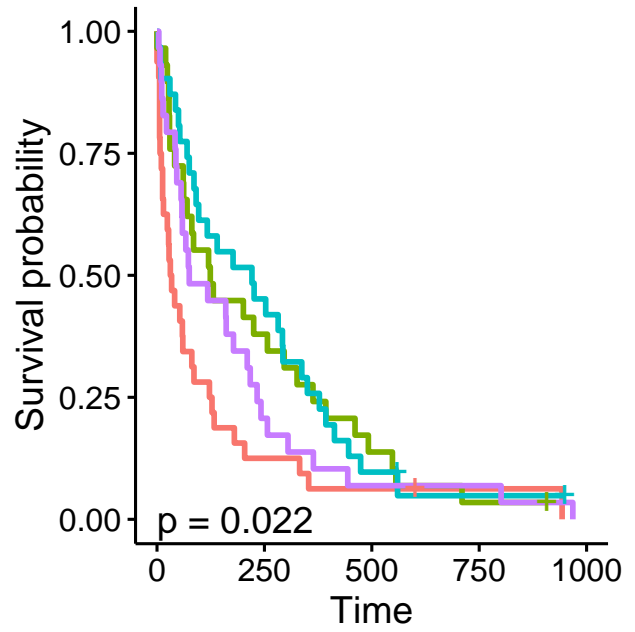
Curative Intent

Strata + Q1 + Q2 + Q3 + Q4



Not Curative Intent

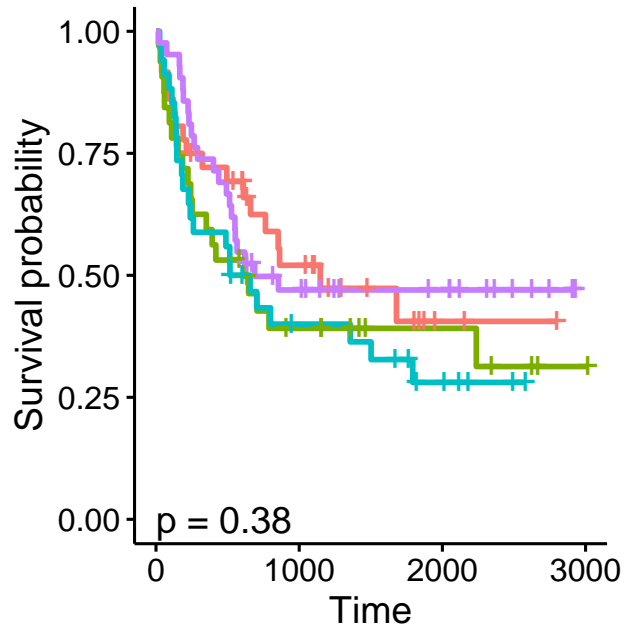
Strata + Q1 + Q2 + Q3 + Q4



CD34neg.CD123pos vs Survival

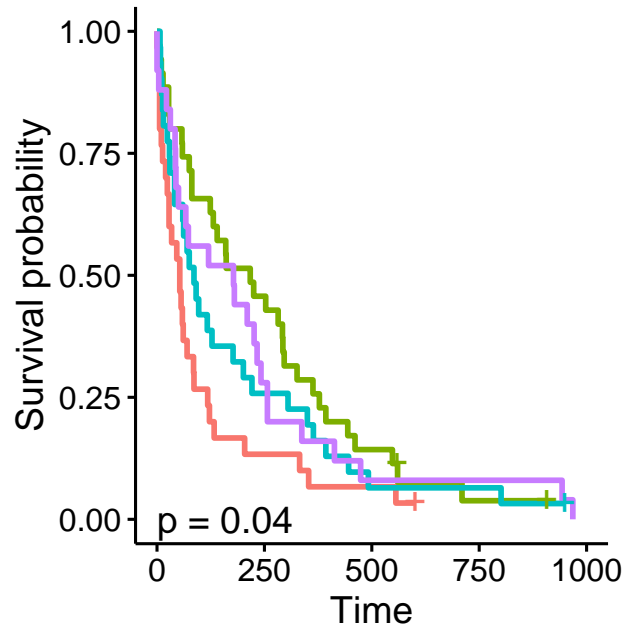
Curative Intent

Strata + Q1 + Q2 + Q3 + Q4



Not Curative Intent

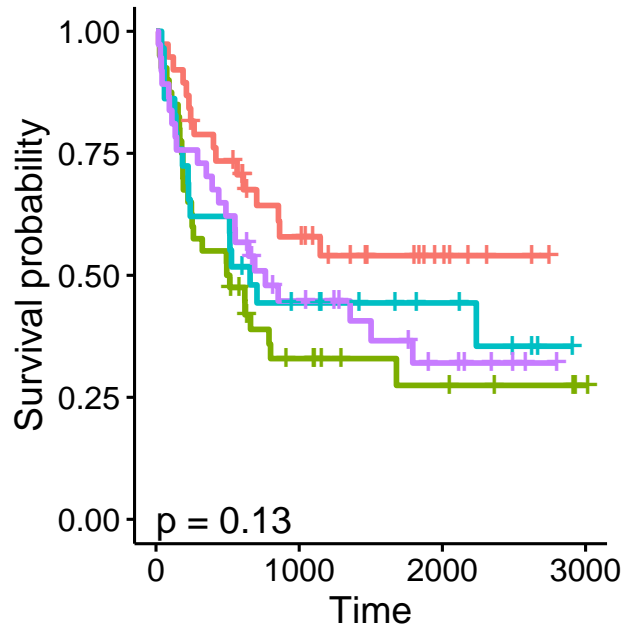
Strata + Q1 + Q2 + Q3 + Q4



CD34pos.CD117pos vs Survival

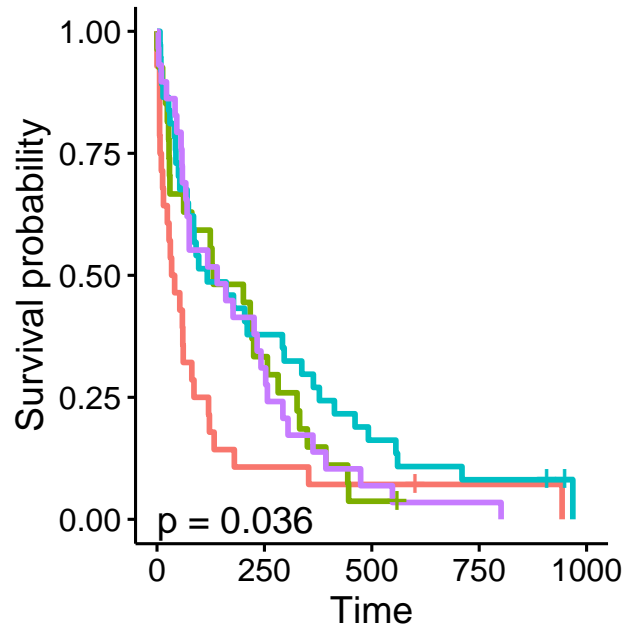
Curative Intent

Strata + Q1 + Q2 + Q3 + Q4



Not Curative Intent

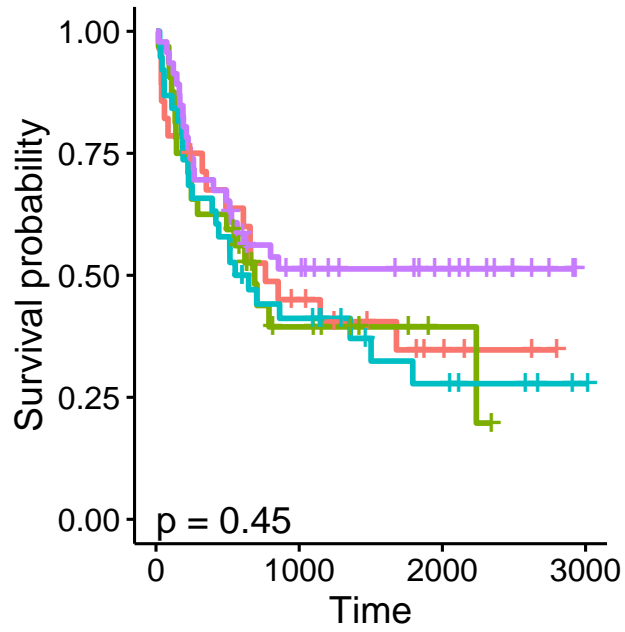
Strata + Q1 + Q2 + Q3 + Q4



CD34neg.CD117pos vs Survival

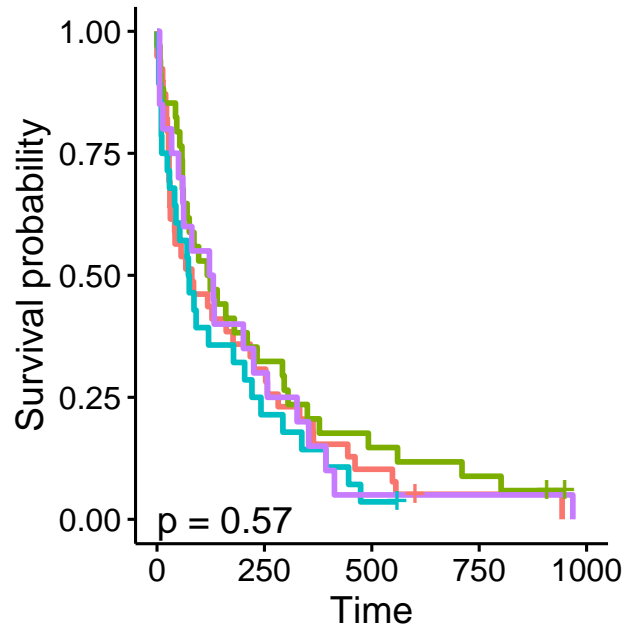
Curative Intent

Strata + Q1 + Q2 + Q3 + Q4



Not Curative Intent

Strata + Q1 + Q2 + Q3 + Q4



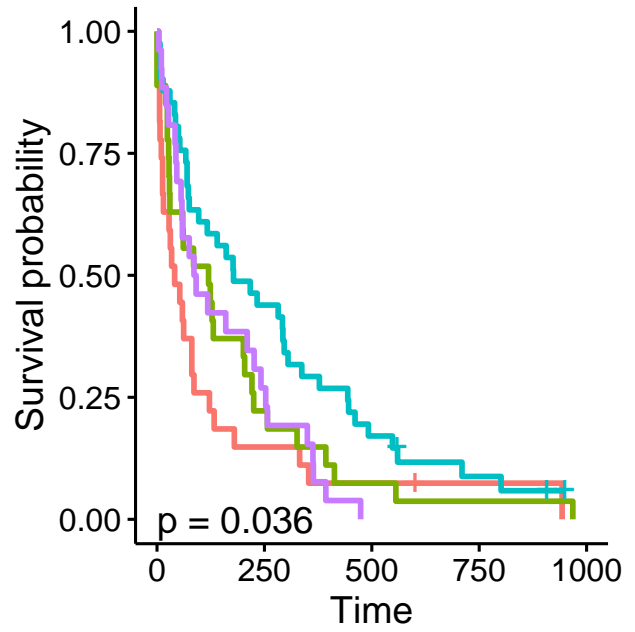
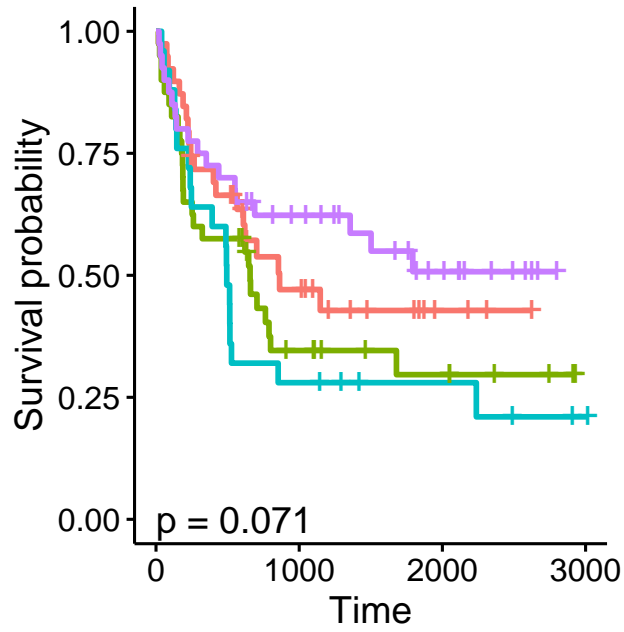
CD34pos.CLEC12Apos vs Survival

Curative Intent

Not Curative Intent

Strata + Q1 + Q2 + Q3 + Q4

Strata + Q1 + Q2 + Q3 + Q4



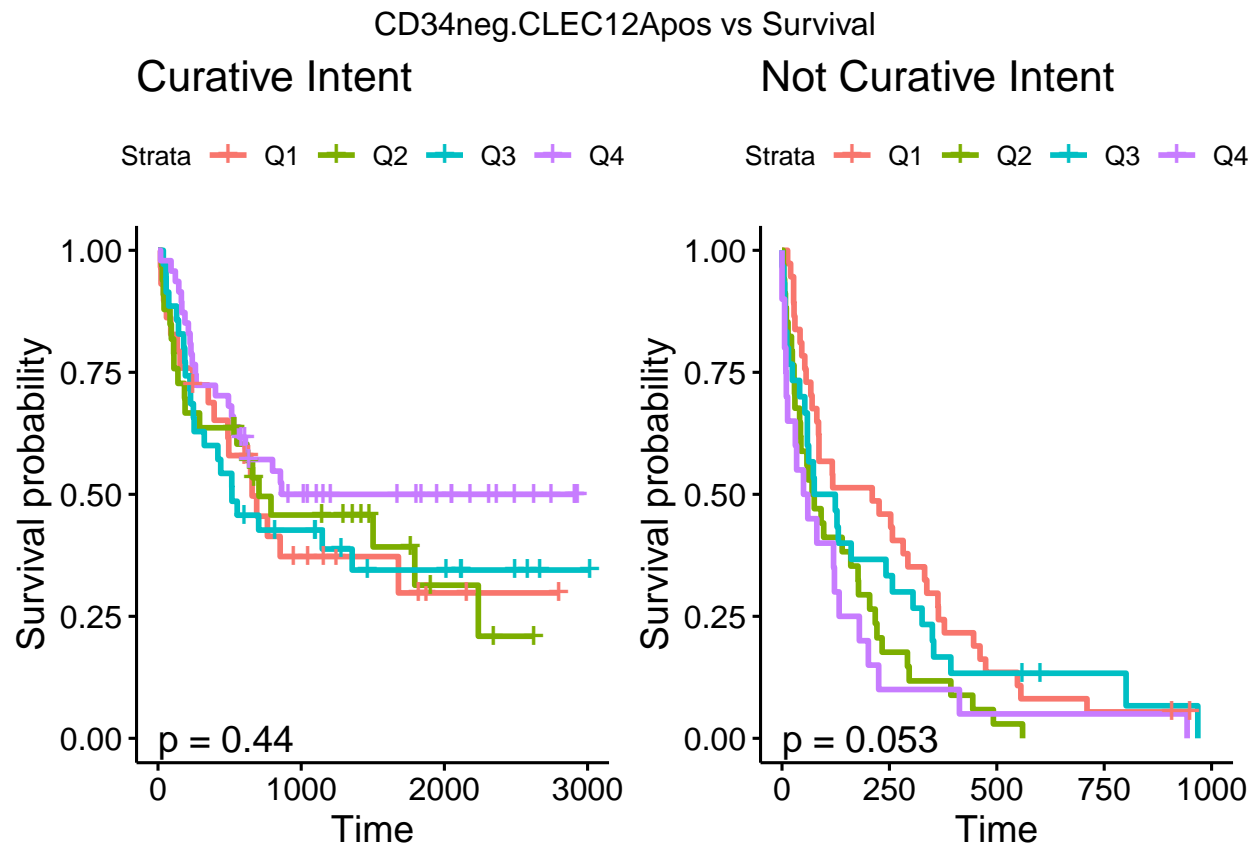


Table 43: Log rank p-values for immunophenotypes split by median or quartiles

	median.curative	median.noncurative	quart.curative	quart.noncurative
CD34pos.frq	0.730	0.017	0.048	0.008
CD117pos.frq	0.466	0.663	0.589	0.120
CD123pos.frq	0.721	0.725	0.240	0.955
CLEC12Apos.frq	0.010	0.196	0.005	0.320
CD34pos.CD123pos.frq	0.236	0.211	0.132	0.022
CD34neg.CD123pos.frq	0.962	0.848	0.376	0.040
CD34pos.CD117pos.frq	0.670	0.048	0.132	0.036
CD34neg.CD117pos.frq	0.536	0.423	0.451	0.575
CD34pos.CLEC12Apos.frq	0.652	0.042	0.071	0.036
CD34neg.CLEC12Apos.frq	0.328	0.577	0.443	0.053