Meta-analysis_results

Alessandra Meddis

```
load("compR.Rdata")
```

Individual patient data meta-analysis for competing events

We present the analysis for a simulated data set characterized by:

- 5012 observations collected from 23 trials
- 4 subgroups of treatment type (adjuvant,concomitant,concomitant+adjuvant, induction)
- 3 competing events: local relapse (failure_type.n=1), distant relapse(failure_type.n=2), death without failure(failure_type.n=3)
- 50% of right censoring (failure_type.n=0)
- age as a continuous variable

```
head(compR)
```

```
##   trial       trt       grpCT patid failure_time          failure_type
## 1     1 Treatment Concomitant  3580    17.820045       Distant failure
## 2     1 Treatment Concomitant  3581    17.982941              No event
## 3     1   Control Concomitant  3582    18.355024              No event
## 4     1   Control Concomitant  3583    11.961514              No event
## 5     1 Treatment Concomitant  3584     1.872994              No event
## 6     1 Treatment Concomitant  3585    11.467748 Loco-regional failure
##     age failure_type.n trt_temp
## 1 41.37              2        1
## 2 41.71              0        1
## 3 29.27              0        0
## 4 49.76              0        0
## 5 44.50              0        1
## 6 75.42              1        1
```

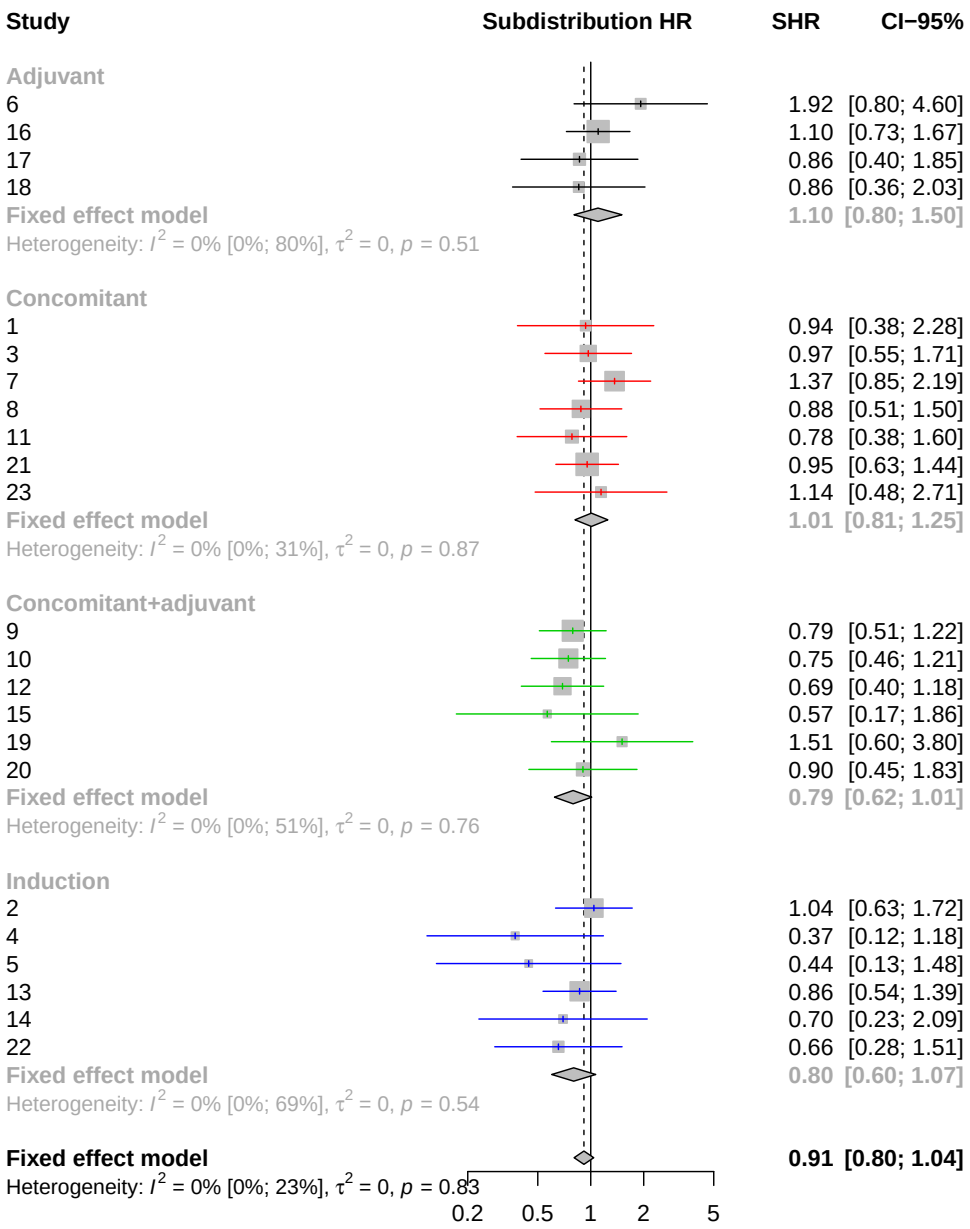
The objective is to analyze this data in order to detect the treatment effect (addition of chemotherapy to radiotherapy). We propose a guideline of statistical methodology that can be applied in this context.

Forest plot

We stratify on the trial and we want to have the information for all the treatment subgroups. We specify that we are interested on local relapse (failcode=1) and also the quantity of interest:

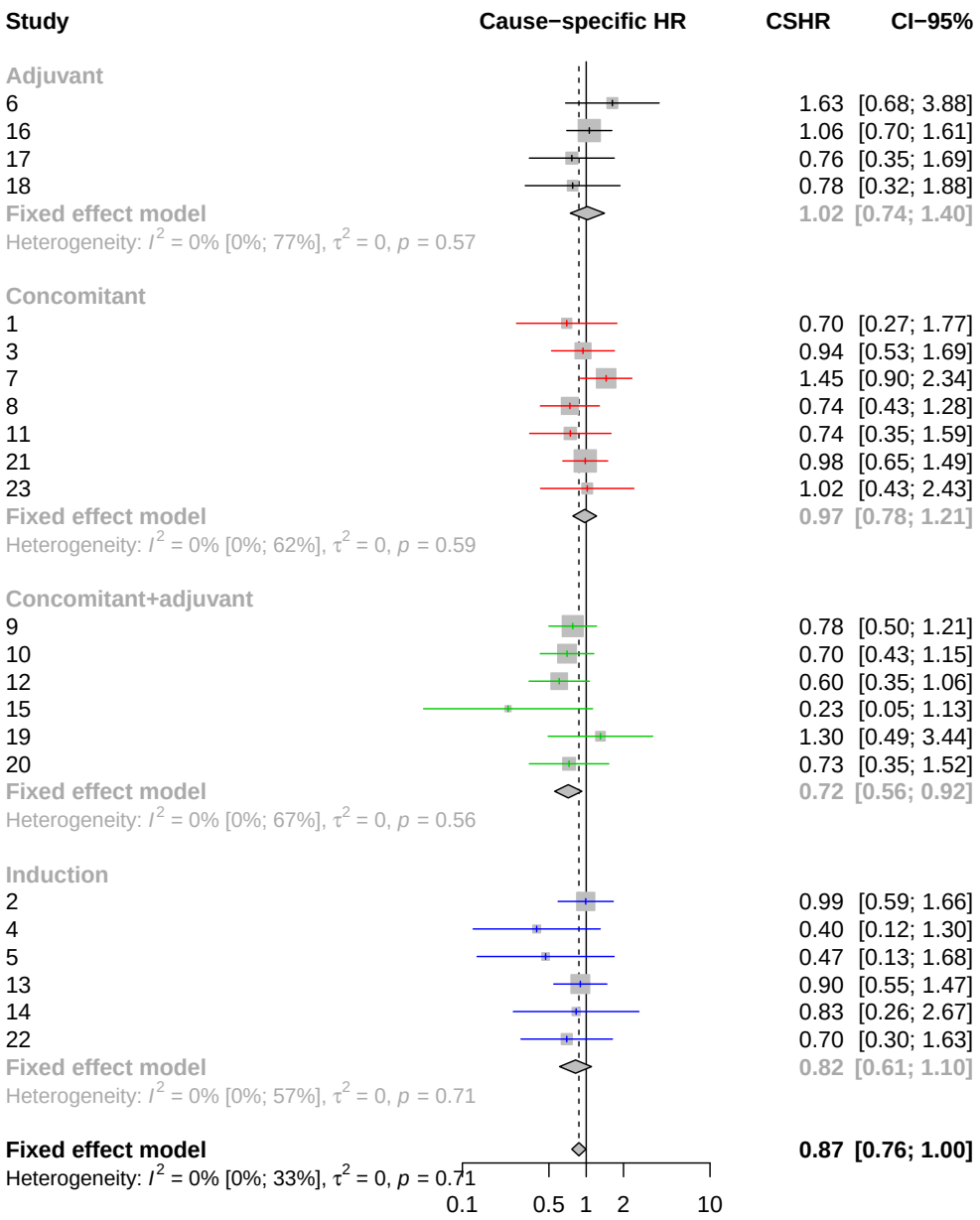
SHR

```
#we call the function described in details n the .R file for localr relapse and SHR
resSHR<-ForestPlot(ftime=compR$failure_time,fstatus=compR$failure_type.n,trt=compR$trt_temp,
           failcode=1, strata=compR$trial,group= compR$grpCT, quantity="SHR")
```



CSHR

```
#we call the function described in details n the .R file for local relapse and CSHR
resCSHR<-ForestPlot(ftime=compR$failure_time,fstatus=compR$failure_type.n,trt=compR$trt_temp,
           failcode=1, strata=compR$trial,group = compR$grpCT, quantity="CSHR")
```



Stratified Fine-Gray model

In an IPD meta-analysis, postulating a common baseline hazard across trials is not tenable due to varying patient populations and treatment. This issue can be addressed via a stratification on the study which relaxes the assumption of proportional hazards regression model. We consider a stratified extension proposed by Zhou et al (2011) implementd in the package crrSC for the estimation of treatment effect (one stage strategy with fixed effect model)

```
library(crrSC)
mod_str=crrs(ftime=compR$failure_time, fstatus=compR$failure_type.n,
             cov1=compR$trt_temp, strata=compR$trial, failcode=1,
             cencode=0)
```

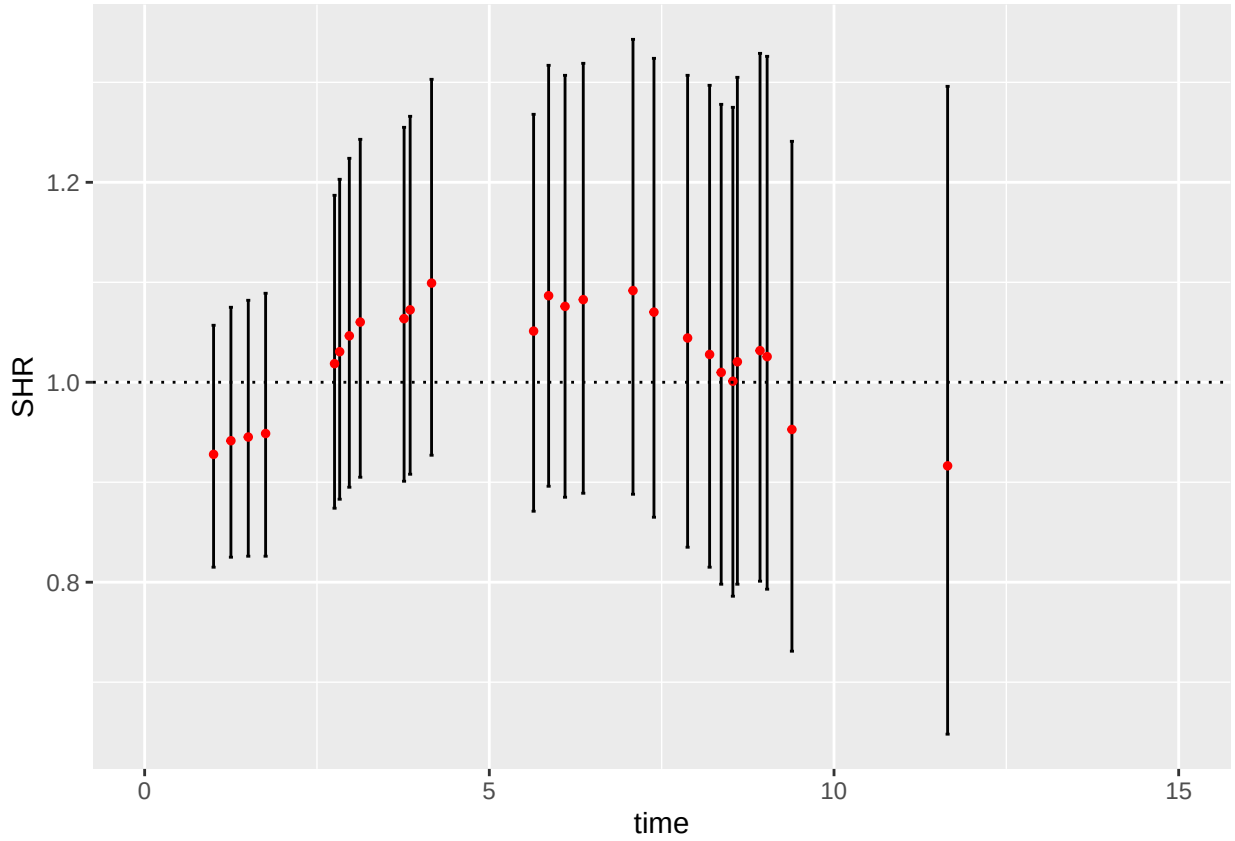
```
SHR_str=c(SHR=exp(mod_str$coef),
            inf=round(exp(mod_str$coef-1.96*sqrt(mod_str$var[1,1])),3),
            sup=round(exp(mod_str$coef+1.96*sqrt(mod_str$var[1,1])),3))

SHR_str
```

```
##       SHR       inf       sup
## 0.9162829 0.8060000 1.0410000
```

Landmark of treatment effect

1.An IPD meta-analysis will typically includes small and large studies with different follow up times (FUP). Landmark for all data considering as landmark times the FUP of the studies (difference in prediction if the follow-up is longer/shorter)

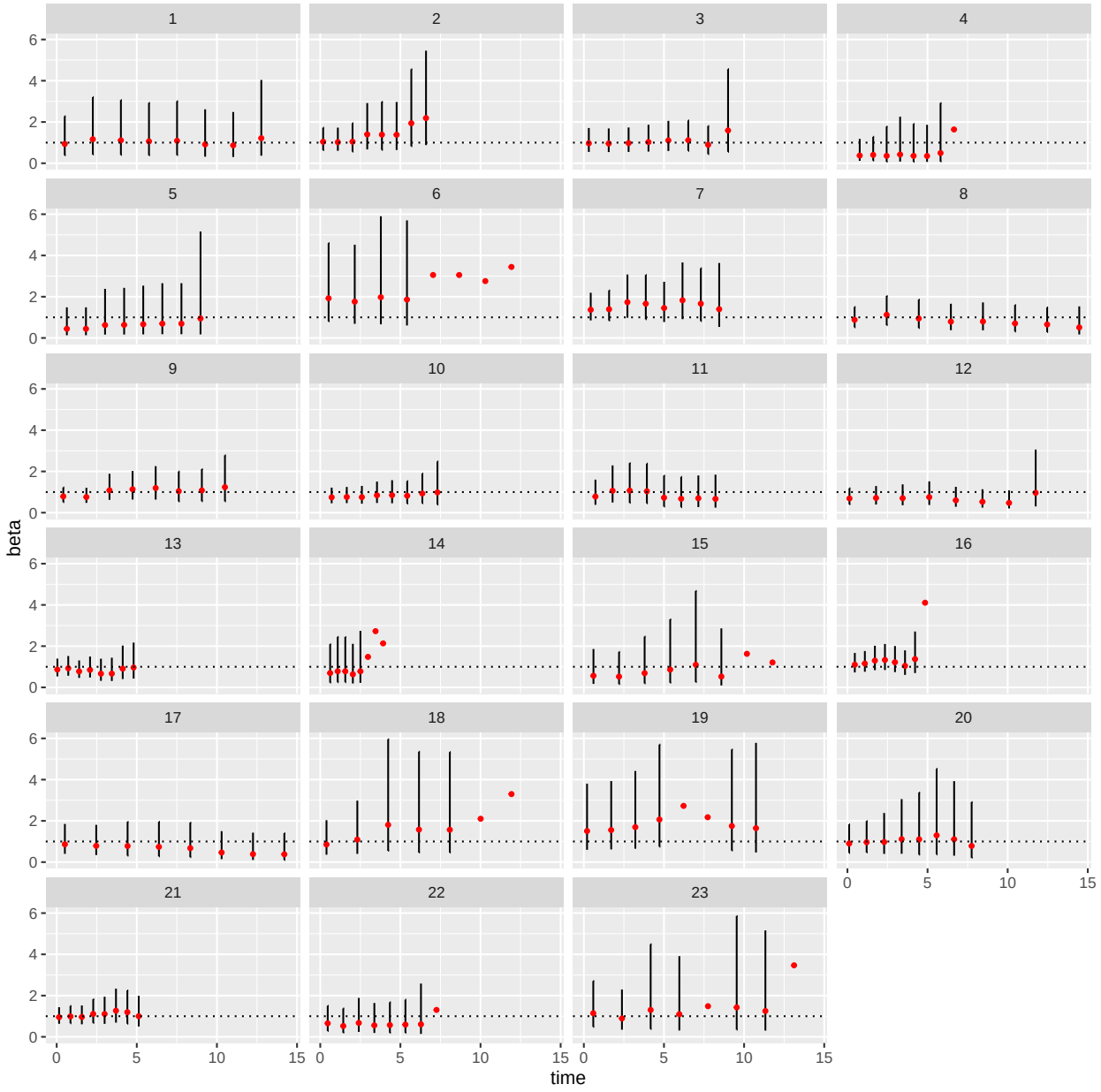
```
library(dplyr)
library(crrSC)
library(ggplot2)


#calculate the FUP for each study
FUP<-tapply(compR$failure_time,compR$trial,median)
#create the landmark times sequence
#before min(FUP) we define more 4 ftimes where to estimate the treatment effect(SHR)
ldmt<-c(seq(1,min(FUP)-1,length.out = 4),sort(FUP))
#estimate of the SHR for each landmark time
SHR_l<-sapply(1:(length(ldmt)-1), FUN=function(k){
  #select subjects at risk for the specific landmark time (ldmt)
  admc<-filter(compR, failure_time>ldmt[k])
  #once we defined the database we use a stratified Fine-Gray model
  mod_str=crrs(ftime=admc$failure_time, fstatus=admc$failure_type.n,
               cov1=admc$trt_temp, failcode=1, strata=admc$trial, #local relapse(failcode=1)
               cencode=0)
  #SHR and c.i.
  SHR_s=exp(mod_str$coef)
  inf=round(exp(mod_str$coef-1.96*sqrt(mod_str$var[1,1])),3)
  sup=round(exp(mod_str$coef+1.96*sqrt(mod_str$var[1,1])),3)
  c(SHR_s,inf,sup)})

#we create the database for the ggplot
SHR_l<-t(SHR_l)
SHR_l<-data.frame(time=ldmt[1:(length(ldmt)-1)], SHR=SHR_l[,1],
                  inf=SHR_l[,2],sup=SHR_l[,3])

ggplot(SHR_l,
       aes(x=time,ymin=inf, lower=inf ,
           middle=SHR, upper=sup, ymax=sup )) +
  geom_errorbar(stat="identity")+
  xlim(0,15)+
  geom_point(aes(x=time,y=SHR), lwd=1, col="red")+
  geom_abline(slope=0,intercept = 1,lty=3)
```



2. Landmark for each study (time-varying trt effect–> non proportionality) we consider the landmark times in order to have enough observations

```
library(dplyr)
library(cmprsk) #SHR in each strata
library(ggplot2)

trial<-unique(compR$trial)
#function foe each trial i
SHRt_trial<-function(i){
  #define the database for the strata i
  temp<-compR[compR$trial==trial[i],]
  #define the landmark time (sequence of 8 times)
  tstar<-seq(min(temp$failure_time),
             quantile(filter(temp,failure_type.n==2)$failure_time,0.7),length.out = 8)
  dit<-length(tstar)
  #for each time we have SHR in this strata (trial) i
  a<-sapply(1:dit, FUN=function(k){
    #subjects at risk
    admc<-filter(temp, failure_time>=tstar[k])
    #Fine-Gray model
    mod_str=crr(ftime=admc$failure_time, fstatus=admc$failure_type.n,
                cov1=admc$trt_temp, failcode=1, #local relapse(failcode=1)
                cencode=0)
    #SHR(tstar) and c.i.
    SHR_s=exp(mod_str$coef)
    inf=round(exp(mod_str$coef-1.96*sqrt(mod_str$var[1,1])),3)
```

```
    sup=round(exp(mod_str$coef+1.96*sqrt(mod_str$var[1,1])),3)
    c(SHR_s,inf,sup)})

  cbind(time=tstar,t(a),trial=rep(i,length(tstar)))
}

#call the function for each trial and we create the database
SHRt_t<-do.call(rbind,lapply(trial,FUN = SHRt_trial))

#prepare data for ggplot
colnames(SHRt_t)=c('time',"beta","inf", "sup","trial")
SHRt_t<-data.frame(SHRt_t)
#plot
ggplot(SHRt_t,
       aes(x=time,ymin=inf, lower=inf ,
           middle=beta, upper=sup, ymax=sup )) +
  geom_errorbar(stat="identity")+
  geom_point(aes(x=time,y=beta), lwd=1, col="red")+
  geom_abline(slope=0,intercept = 1,lty=3)+
  ylim(0,6)+
  facet_wrap( ~ trial, ncol=4)
```



Proportionality of SHR

1.To check for the assumption of the proportionality of subdistribution hazards ratio we implement the PSH.test proposed by Zhou et al (2013) for three different time functions $t^2, t, log(t)$.

```
#numeric variable for treatment
compR$trt_temp=ifelse(compR$trt=="Control", 0,1)

#goodness of fit for each stratum (trial) id
gof<-function(id){
  #considero il trial id
  temp<-compR[compR$trial==id,]
  #f(t)=t^2
```

```
  gofx2<-psh.test(time=temp$failure_time, fstatus=temp$failure_type.n, z=temp$trt_temp,
                  D=c(1),tf=function(x) c(x^2))
  #f(t)=t
  gofx<-psh.test(time=temp$failure_time, fstatus=temp$failure_type.n, z=temp$trt_temp,
                 D=c(1),tf=function(x) c(x))
  #f(t)=log(t)
  goflog<-psh.test(time=temp$failure_time, fstatus=temp$failure_type.n, z=temp$trt_temp,
                   D=c(1),tf=function(x) c(log(x)))

  #pvalue for each test
  c(id,gofx2[1,5],gofx[1,5],goflog[1,5])
}

trial<-unique(compR$trial)
#data with trial id and pvalue for each test
gof_trial<-do.call(rbind,lapply(trial,FUN = gof))
#which trial reject (pvalue,0.05) the null hypothesis for t^2 (non proportionality)
np_indx2<-list(trial=which(gof_trial[,2]<0.05),pvalue=gof_trial[which(gof_trial[,2]<0.05),2])
#which trial reject (pvalue,0.05) the null hypothesis for t (non proportionality)
np_indx<-list(trial=which(gof_trial[,3]<0.05),pvalue=gof_trial[which(gof_trial[,3]<0.05),3])
#which trial reject (pvalue,0.05) the null hypothesis for log(t) (non proportionality)
np_indlogx<-list(trial=which(gof_trial[,4]<0.05),pvalue=gof_trial[which(gof_trial[,4]<0.05),4])

#Results for the data (local relapse)

np_indx2
```

```
## $trial
## [1] 18
##
## $pvalue
## p-value
##  0.0172
```

```
np_indx
```

```
## $trial
## [1] 18
##
## $pvalue
## p-value
##  0.0173
```

```
np_indlogx
```

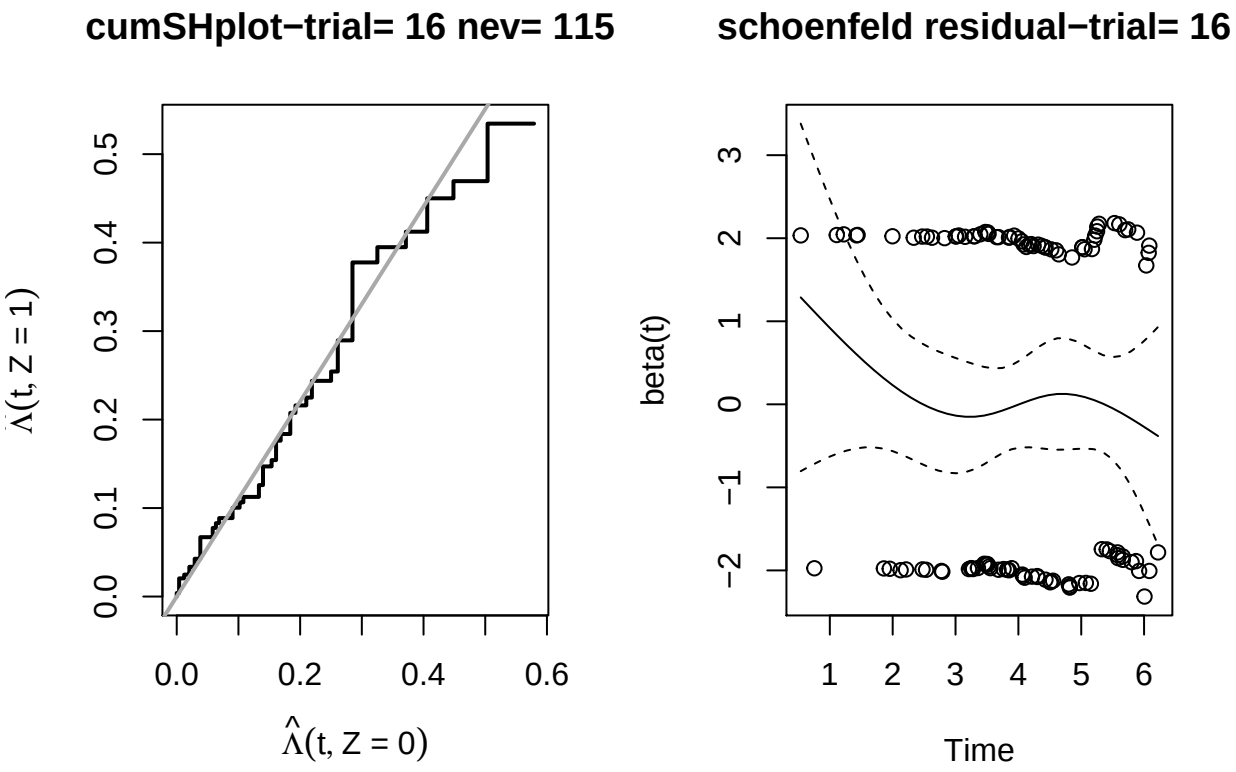
```
## $trial
## [1] 18
##
## $pvalue
## p-value
##  0.0234
```

2. Graphical methods to look for non proportionality are the Schoenfeld's resilduals plot and the cumulative subdistribution functions plot. Plotting the Schoenfeld residuals, we would expect mean equal to 0 across time and, a non constant average is related to misspecification of the PSH. Similarly, for binary covariate, the cumulative subdistribution hazards plot we would expect a constant ratio in case of PSH.

We present the result for trial 18 which rejected the null hypotheis of the PSH.test previously employed:

```
##For the Cumulative subdistribution hazards we estimate the cumulative incidence function
#considering the specific case of a multi-state model (A-J model)
#where the only possible transitions are initial state - event with event=local,distant or death

library(etm)

## create the etm database
compR$from<-rep(8,nrow(compR)) #we define 8 as initial state
compR$id<-compR$patid
compR$time<-compR$failure_time
compR$to<-compR$failure_type.n

#define the transition matrix
tra<-matrix(FALSE,4,4)
#put TRUE to 1-->2,3,4 because we are in competing risk settings
#and in 1 is the initial state and 2,3,4 the possible events
tra[1,2:4]<-TRUE

#for each trial id we estimate the cumulative SHR (distant relapse)
cumSH<-function(id){
  #data for each trial
  temp<-compR[compR$trial==id,]
  #A-J model for treat 1 or 0
  #the possible state are 8,1,2,3 (initial state,local, distant, death)
  AJ.z1<-etm(temp[temp$trt_temp == 1, ], c("8","1","2","3"), tra, "0", 0)
  AJ.z0<-etm(temp[temp$trt_temp == 0, ], c("8","1","2","3"), tra, "0", 0)
  ###time of events (for both treated and control)
  times <- sort(c(AJ.z0$time, AJ.z1$time))
  #estimation of probability of transition (form initial state to distant relapse)
  cif.z0 <- trprob(AJ.z0, tr.choice = "8 2", timepoints = times)
  cif.z1 <- trprob(AJ.z1, tr.choice = "8 2", timepoints = times)
  #cum sub hazards from the cumulative incidence function
  sub.haz.z0 <- cumsum(1 -((1 - cif.z0) /
                             (1 - c(0, cif.z0[-length(cif.z0)]))))
  sub.haz.z1 <- cumsum(1 -((1 - cif.z1) /
                             (1 - c(0, cif.z1[-length(cif.z1)]))))

  data.frame(rep(id,length(times)),sub.haz.z0,sub.haz.z1,times)
}

cumSH_data<-do.call(rbind,lapply(trial,FUN = cumSH))
colnames(cumSH_data)<-c("trial","cumSH0","cumSH1","times")

# plot cumSH and Schoenfeld residuals
#trial 16
par(mfrow=c(1,2))
  ##cumSH
  tmp<-cumSH_data[cumSH_data$trial==16,]
  temp<-compR[compR$trial==16,] #for the SCh plot
  plot(tmp$cumSH0, tmp$cumSH1, lwd = 2, type = "s",
       xlab = expression(hat(Lambda)(t, "Z = 0")),
       ylab = expression(hat(Lambda)(t, "Z = 1")),
```

```
      main=paste("cumSHplot-trial=",16,"nev=",nrow(filter(temp,to==2)),""))

  shfit.lr <- with(compR[compR$trial==16,], crr(time, failure_type.n ,trt_temp ,
                                                failcode = 1, cencode = 0))
  abline(a = 0, b = exp(shfit.lr$coef), col = "darkgray", lwd = 2)

  #Sch residuals plot
  cox.rel.mv <- coxph(Surv(failure_time, failure_type.n == 2) ~ trt_temp, temp)
  plot(cox.zph(cox.rel.mv, transform='identity'), ylab="beta(t)",
       main=paste("schoenfeld residual-trial=",16,""))
```



Detection of possible interactions

IPD meta-anlaysis proide some individual-level covariate that allow us to check for interactions with the treatment to identify the subgroup of patients that benefit most fromt he treatment. In this case, we are interested in a possible interaction between treatment and age.

```
 ##design matrix
M<-model.matrix( ~ 0+ compR$trt +compR$age + compR$trt:compR$age )

#stratified Fine-Gray model
mod_str=crrs(ftime=compR$failure_time, fstatus=compR$failure_type.n,
             cov1=M[,2:4], strata=compR$trial, failcode=1,
             cencode=0)
#beta_trt<-c(coef=round(exp(mod_str$coef[1]),3),
```

```
#             inf=round(exp(mod_str$coef[1]-1.96*sqrt(mod_str$var[1,1])),3),
#             sup=round(exp(mod_str$coef[1]+1.96*sqrt(mod_str$var[1,1])),3))

#interaction term
beta_int<-c(coef=round(exp(mod_str$coef[3]),3),
            inf=round(exp(mod_str$coef[3]-1.96*sqrt(mod_str$var[3,3])),3),
            sup=round(exp(mod_str$coef[3]+1.96*sqrt(mod_str$var[3,3])),3))

beta_int
```

```
##  coef   inf   sup
## 0.991 0.982 0.999
```

Interaction age:treatment modality

we want to understant if the age is significant in the treatment efficacy for a specific treatment type thus we first stratify on the treatment type (e.g. adjuvant) and we employ the stratified (on trials) Fine-Gray model in this subgroup With the treatment variable and the interaction term.

```
#first stratification (treatment modality = adjuvant)
adj<-filter(compR,grpCT=="Adjuvant")
#design Matrix
M2<-model.matrix( ~ 0+ adj$trt +adj$age + adj$trt:adj$age )
#stratified Fine-Gray model
mod_adj<-crrs(ftime=adj$failure_time, fstatus=adj$failure_type.n,
              cov1=M2[,2:4],
              strata=adj$trial, failcode=1,
              cencode=0)
#interaction trt:age for adjuvant chemotherapy
beta_adj_int<-c(coef=round(exp(mod_adj$coef[3]),3),
            inf=round(exp(mod_adj$coef[3]-1.96*sqrt(mod_adj$var[3,3])),3),
            sup=round(exp(mod_adj$coef[3]+1.96*sqrt(mod_adj$var[3,3])),3))

beta_adj_int
```

```
##  coef   inf   sup
## 1.011 0.991 1.031
```