

Vignette for R package PBIR

Bo Huang, Xiaodong Luo, and Lu Tian

May 6, 2020

Introduction

The PBIR package contains several key functions for estimating and comparing the probability of being in response (PBIR), and for estimating and comparing the mean duration of response (DOR) in ITT population (Tsai et al. 2017 and Huang et al. 2020). Let T_1 and T_2 denote the time to response and time to disease progression or death, respectively. We assume that $T_1 = \infty$, if disease progression or death occurs first. The PBIR at time t is defined as the probability that a patient of being in response (as a responder and has not progressed or died) at time point t , i.e., $T_1 \leq t$ and $t < T_2$. Let $T_3 = \min(T_1, T_2)$, then

$$PBIR(t) = P(T_2 > t) - P(T_3 > t),$$

which can be estimated as the difference between two Kaplan Meier estimates for the survival function of T_2 and T_3 , respectively, in the presence of censoring. Since the survival functions of T_2 or T_3 may not be estimable over their respective entire support due to right censoring, PBIR is in general only identifiable within a time window $[0, \tau]$, where τ is a time point no more than the longest followup time in the data (time to response, progression, or death, whichever is earlier). Furthermore, since these two Kaplan Meier estimates are correlated, the variance estimates of the estimated PBIRs need to account for this correlation appropriately.

The PBIR package allows researchers to calculate the PBIR curve over a time window $[0, \tau]$ and compare PBIR curves between two groups. This document provides illustrative examples of how to use R functions from “PBIR” package to make statistical inferences on PBIR in different settings. First, we load the “PBIR” and “survival” packages

```
library(PBIR)
library(survival)
```

Estimating PBIR over a given time window

In this section, we will show how to use the R function “PBIR1” to estimate PBIR over a time window. To this end, we first need to generate data using following R code:

```
set.seed(100)
n=100
error=rnorm(n)
tr=exp(rnorm(n)+error+0.5)
tp=exp(rnorm(n)+error)
tr[tp<tr]=Inf
tc=runif(n, 3, 8.5)
t2response=pmin(tr, tc)
delta_response=1*(tr<tc)
t2progression=pmin(tp, tc)
delta_progression=1*(tp<tc)
```

In this simulation, the time to treatment response and time to disease progression, i.e. $(T_1, T_2)'$, are generated from a bivariate log-normal distribution

$$\exp\left(N\left\{\begin{pmatrix} 0.5 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\right\}\right)$$

except that we set $T_1 = \infty$ if $T_1 > T_2$. A common censoring time C is generated from a uniform distribution $U(3, 8.5)$. The generated data consist of n copies of $(X_1, \delta_1, X_2, \delta_2)$, where $(X_j, \delta_j) = \{\min(T_j, C), I(T_j < C)\}$, $j = 1, 2$. The first few simulated observations are

```
round(head(cbind(t2response, delta_response=1*(tr<tc), t2progression, delta_progression)), 4)
```

```
##      t2response delta_response t2progression delta_progression
## [1,]      3.1747              0           0.6225              1
## [2,]      6.1884              0           0.7984              1
## [3,]      0.9531              1           2.1678              1
## [4,]      3.6890              0           3.6890              0
## [5,]      0.4313              1           3.1117              1
## [6,]      8.0234              0           0.4952              1
```

Then we can estimate the PBIR based on generated data. The R code is

```
fit=PBIR1(t2PROGRESSION=t2progression, STATUS_PROGRESSION=delta_progression,
          t2RESPONSE=t2response, STATUS_RESPONSE=delta_response)
```

Here, “t2PROGRESSION” is time to disease progression, death, or censoring, whichever is earlier; “STATUS_PROGRESSION” is the binary indicator for if death/progression occurs earlier than censoring; “t2RESPONSE” is time to treatment response or censoring, whichever is earlier; and “STATUS_RESPONSE” is the binary indicator for if treatment response occurs earlier than censoring. We didn’t specify any value for “time”, so that PBIR will be estimated over the largest time window in which it is identifiable. We may display the estimated PBIRs at some selected time points as

```
round(fit[95:100, ], 4)
```

```
##      time PBIR      std ci.low ci.up
## 95  1.9546 0.07 0.0248 0.0345 0.1370
## 96  2.0311 0.08 0.0263 0.0413 0.1492
## 97  2.0963 0.08 0.0263 0.0413 0.1492
## 98  2.1209 0.08 0.0263 0.0413 0.1492
## 99  2.1210 0.08 0.0263 0.0413 0.1492
## 100 2.1436 0.09 0.0277 0.0485 0.1611
```

The first column represents time points at which PBIR is estimated, the second column represents the estimated PBIR, the third column represents the standard error of the estimated PBIR, and the last two columns represent the 95% point-wise confidence interval of the PBIR. For example, the estimated PBIR at $t = 2.03$ is 0.08 (95% confidence interval: 0.041 to 0.149). Furthermore, we may plot the estimated PBIR curve along with its 95% point-wise confidence interval

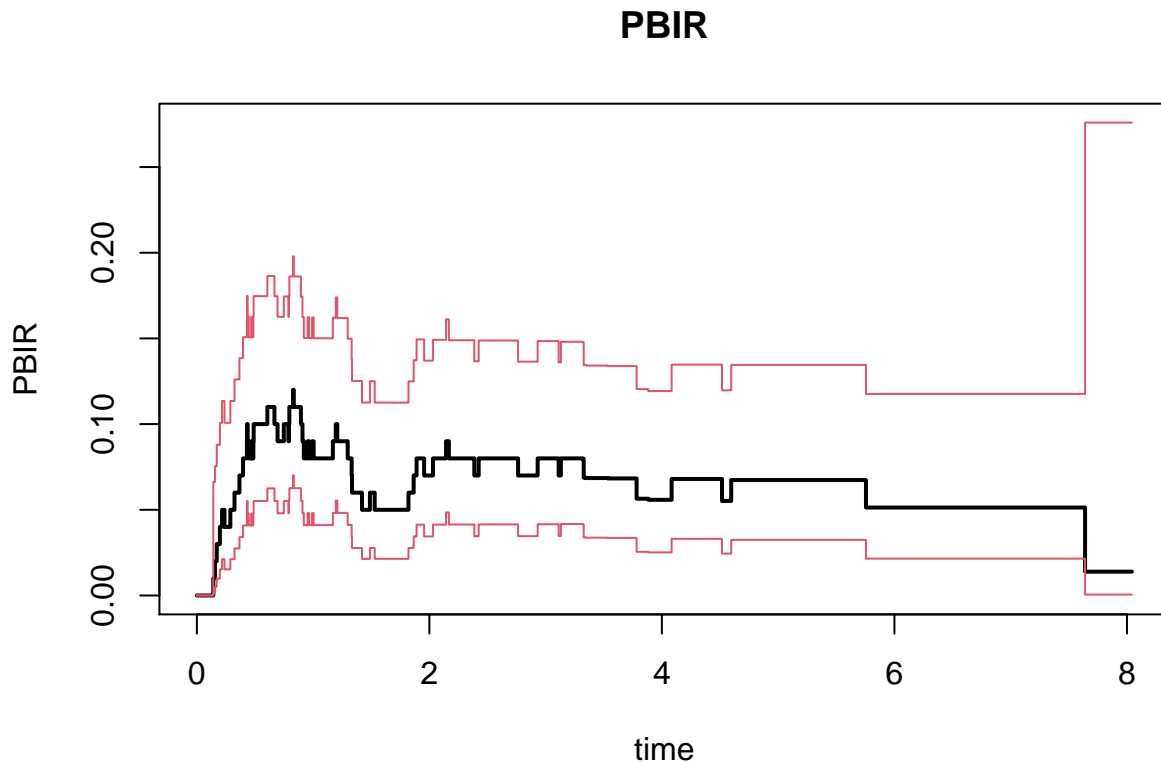
```
tt=c(0, fit$time)
PBIR=c(0, fit$PBIR)
ci.up=c(0, fit$ci.up)
ci.low=c(0, fit$ci.low)

B=length(tt)
tt=rep(tt, rep(2, B))[-1]
PBIR=rep(PBIR, rep(2, B))[-(2*B)]
ci.up=rep(ci.up, rep(2, B))[-(2*B)]
ci.low=rep(ci.low, rep(2, B))[-(2*B)]
```

```

plot(range(tt), range(na.omit(c(ci.low, ci.up))),
     xlab="time", ylab="PBIR",
     main="PBIR", type="n")
lines(tt, PBIR, col=1, lwd=2)
lines(tt, ci.low, col=2)
lines(tt, ci.up, col=2)

```



We also can estimate the PBIR at selected time points. The R code for estimating PBIR at $t = 2, 4$ and 6 is

```

fit=PBIR1(t2PROGRESSION=t2progression, STATUS_PROGRESSION=delta_progression,
          t2RESPONSE=t2response,      STATUS_RESPONSE=delta_response,
          time=c(2,4,6))

round(fit, 4)

```

```

## time PBIR std ci.low ci.up
## 1 2 0.0700 0.0248 0.0345 0.1370
## 2 4 0.0558 0.0223 0.0251 0.1193
## 3 6 0.0513 0.0224 0.0215 0.1176

```

If we want to estimate the PBIR at $t = 7$ and 10 , then the R code is

```

fit=PBIR1(t2PROGRESSION=t2progression, STATUS_PROGRESSION=delta_progression,
          t2RESPONSE=t2response,      STATUS_RESPONSE=delta_response,
          time=c(7,10))

```

```

## The PBIR is not identifiable at some selected time points,

```

```
## which are replaced by the maximum time point, where it is identifiable.
```

```
round(fit, 4)
```

```
##      time   PBIR   std ci.low ci.up
## 1 7.0000 0.0513 0.0224 0.0215 0.1176
## 2 8.0431 0.0139 0.0231 0.0005 0.2760
```

The output is PBIR at $t = 7$ and $t = 8.043$ rather than $t = 7$ and $t = 10$, because PBIR at $t = 10$ is not estimable and $t = 10$ is replaced by $t = 8.043$.

Two Group Comparisons

In this section, we will show how to the R function “PBIR2” to compare PBIR over a time window between two groups. To this end, we first simulate data from two groups using the following R code:

```
set.seed(100)
n=200
trt=rbinom(n, 1, 0.5)
error=rnorm(n)
tr=exp(rnorm(n)+error-trt*0.5+0.5)
tp=exp(rnorm(n)+error+trt*0.25)
tr[tp<tr]=Inf
tc=runif(n, 3, 8.5)
t2response=pmin(tr, tc)
delta_response=1*(tr<tc)
t2progression=pmin(tp, tc)
delta_progression=1*(tp<tc)
```

In this simulation, the treatment indicator, i.e., trt , is simulated from a Bernoulli distribution $B(0.5)$, and the time to treatment response and time to disease progression in a given treatment group, i.e., $(T_1, T_2)'|R$, are generated from a bivariate log-normal distribution

$$\exp\left(N\left\{\begin{pmatrix} 0.5 - 0.5R \\ 0.25R \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\right\}\right), R = 0, 1.$$

The first few simulated observations are

```
round(head(cbind(t2response, delta_response=1*(tr<tc), t2progression, delta_progression, trt)), 4)
```

```
##      t2response delta_response t2progression delta_progression trt
## [1,]      6.8168             0           0.1688             1  0
## [2,]      2.3269             1           4.1125             0  0
## [3,]      0.2200             1           0.5701             1  1
## [4,]      7.4737             0           0.3367             1  0
## [5,]      8.0389             0           0.2967             1  0
## [6,]      8.2051             0           0.4662             1  0
```

Then we can estimate the difference in PBIR between two groups (over the largest time window, in which the PBIR curves from two arms are both identifiable) using the R function “PBIR2”:

```
fit2=PBIR2(t2PROGRESSION=t2progression, STATUS_PROGRESSION=delta_progression,
           t2RESPONSE=t2response,          STATUS_RESPONSE=delta_response,
           TRT=trt)
```

The inputs of function “PBIR2” are identical to those of “PBIR1” except that “TRT” is the binary indicator for treatment assignment. The outputs of the function is the difference in PBIRs between two groups. We may display the estimated group difference in PBIRs at selected time points

```
round(fit2[215:220, ], 4)
```

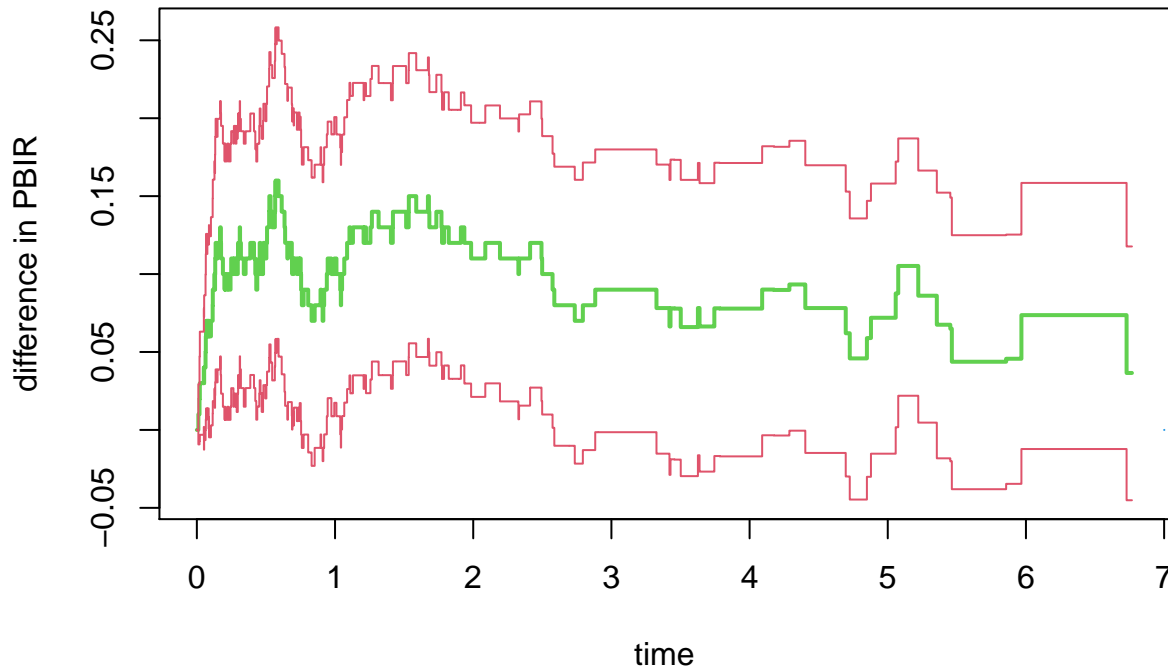
```
##      time diff.in.PRIB      std ci.low ci.up
## 215 1.9358          0.12 0.0457 0.0298 0.2083
## 216 1.9435          0.12 0.0457 0.0298 0.2083
## 217 1.9849          0.11 0.0450 0.0212 0.1971
## 218 2.0054          0.11 0.0450 0.0212 0.1971
## 219 2.0469          0.11 0.0450 0.0212 0.1971
## 220 2.0495          0.11 0.0450 0.0212 0.1971
```

The first column of the output represents time points at which PBIR is compared, the second column represents the estimated between-group difference in PBIR, the third column represents the standard error of the estimated difference, and the last two columns represent the 95% point-wise confidence interval of the difference in PBIR. For example, the estimated difference in PBIR between two groups (group 1 vs. group 0) is 0.11 (95% confidence interval: 0.021 to 0.197). We may plot the estimated group difference in PBIR curve along with its 95% point-wise confidence interval.

```
tt=fit2$time
diff=fit2$diff
low=fit2$ci.low
up=fit2$ci.up
B=length(tt)+1

tt=c(0, tt)
diff=c(0, diff)
low=c(0, low)
up=c(0, up)
tt=rep(tt, rep(2, B))[-1]
diff=rep(diff, rep(2, B))[-(2*B)]
low=rep(low, rep(2, B))[-(2*B)]
up=rep(up, rep(2, B))[-(2*B)]

plot(range(c(tt, 0)), range(c(low, up)),
      xlab="time", ylab="difference in PBIR",
      lwd=2, type="n")
lines(tt, diff, lwd=2, col=3)
lines(tt, low, col=2)
lines(tt, up, col=2)
lines(range(fit$time), rep(0, 2), col=4, lty=4)
```



Based on the plot, it is clear that the PBIR at group 1 is statistically significantly higher than PBIR at group 0 around $t = 2$, since the corresponding 95% confidence intervals exclude zero.

Similarly, one may also compare PBIR at user-specified time points. For example, one can use the following code to compare PBIR at $t = 2, 4$, and 6

```
fit2=PBIR2(t2PROGRESSION=t2progression, STATUS_PROGRESSION=delta_progression,
           t2RESPONSE=t2response, STATUS_RESPONSE=delta_response,
           TRT=trt, time=c(2, 4, 6))

round(fit2, 4)
```

```
## time diff.in.PBIR std ci.low ci.up
## 1 2 0.1100 0.0450 0.0212 0.1971
## 2 4 0.0778 0.0482 -0.0170 0.1713
## 3 6 0.0737 0.0437 -0.0123 0.1585
```

If a time point too big such as $t = 10$ is selected, The “PBIR2” function will automatically replace it by the maximum time point at which PBIRs in both groups are estimable and thus comparable:

```
fit2=PBIR2(t2PROGRESSION=t2progression, STATUS_PROGRESSION=delta_progression,
           t2RESPONSE=t2response, STATUS_RESPONSE=delta_response,
           TRT=trt, time=c(2, 4, 6, 10))
```

```
## The PBIR is not identifiable at some selected time points,
## which are replaced by the maximum time point, where it is identifiable.
## The PBIR is not identifiable at some selected time points,
## which are replaced by the maximum time point, where it is identifiable.
```

```
## The PBIR difference is only estimated at time points, where it is identifiable.
```

```
round(fit2, 4)
```

```
##      time diff.in.PRIB      std ci.low ci.up
## 1 2.0000      0.1100 0.0450 0.0212 0.1971
## 2 4.0000      0.0778 0.0482 -0.0170 0.1713
## 3 6.0000      0.0737 0.0437 -0.0123 0.1585
## 4 6.7671      0.0365 0.0416 -0.0451 0.1177
```

Inferences on the Mean Duration of Response

The mean duration of response is defined as $E\{(T_2 - T_1)I(T_2 \geq T_1)\} = E(T_2 - T_3)$, which is the expected duration a patient of being in response, i.e, being a responder (Huang et al. 2018). Note that the duration of response is defined for all patients: in particular, the duration of response is zero for a patient, who has never responded to the treatment. The mean duration of response may not be identifiable beyond the maximum follow-up time. Instead, we are interested in estimating the mean duration of response restricted within a time window $[0, \tau]$:

$$E[\min(T_2, \tau) - \min(T_3, \tau)],$$

which can also be expressed as the area under the PBIR curve

$$\int_0^\tau PBIR(t)dt.$$

Therefore, the mean duration of response can be estimated as

$$\int_0^\tau \widehat{PBIR}(t)dt.$$

Next, we will show how to use the R function “mduration” to estimate the mean duration of response over a given time window. To this end, we first simulate data using the same R code above to generate time to treatment response and time to disease progression from two groups of patients:

```
set.seed(100)
n=200
trt=rbinom(n, 1, 0.5)
error=rnorm(n)
tr=exp(rnorm(n)+error-trt*0.5+0.5)
tp=exp(rnorm(n)+error+trt*0.25)
tr[tp<tr]=Inf
tc=runif(n, 3, 8.5)
t2response=pmin(tr, tc)
delta_response=1*(tr<tc)
t2progression=pmin(tp, tc)
delta_progression=1*(tp<tc)
```

We can estimate the mean duration of response up to $t = 8$ by treatment groups based generated data. The R function “mduration” can be used to achieve this objective:

```
fit1=mduration(t2PROGRESSION=t2progression[trt==1],
              STATUS_PROGRESSION=delta_progression[trt==1],
              t2RESPONSE=t2response[trt==1],
              STATUS_RESPONSE=delta_response[trt==1])

fit0=mduration(t2PROGRESSION=t2progression[trt==0],
              STATUS_PROGRESSION=delta_progression[trt==0],
```

```
t2RESPONSE=t2response[trt==0],
STATUS_RESPONSE=delta_response[trt==0])
```

The outputs consisting of point estimator of mean duration of response and its standard error are

```
fit1
```

```
## $meandor.est
## [1] 1.096093
##
## $meandor.se
## [1] 0.1666711
##
## $time.truncation
## [1] 6.767068
```

```
fit0
```

```
## $meandor.est
## [1] 0.5494218
##
## $meandor.se
## [1] 0.1438812
##
## $time.truncation
## [1] 7.974405
```

The input of the function “mduration” is the same as that of the function “PBIR1”. The output includes a point estimator of mean DOR and its standard error. In this example, the mean duration of response within the time window [0, 6.76] is 1.096(standard error: 0.167) in group 1 and the mean duration of response within the time window [0, 7.97] is 0.549 (0.144) in group 0. If we want to compare mean duration of response between two groups, we need to specify a common time window. Suppose that we want to compare the mean duration of response within the time window [0, 6.75], we can easily compute the Wald-test statistics and the corresponding p-value from outputs of “mduration”:

```
fit1=mduration(t2PROGRESSION=t2progression[trt==1],
              STATUS_PROGRESSION=delta_progression[trt==1],
              t2RESPONSE=t2response[trt==1],
              STATUS_RESPONSE=delta_response[trt==1],
              time.max=6.75)

fit0=mduration(t2PROGRESSION=t2progression[trt==0],
              STATUS_PROGRESSION=delta_progression[trt==0],
              t2RESPONSE=t2response[trt==0],
              STATUS_RESPONSE=delta_response[trt==0],
              time.max=6.75)

diff=(fit1$meandor.est-fit0$meandor.est)
se=sqrt(fit1$meandor.se^2+fit0$meandor.se^2)
z=diff/se
ci=diff+c(-1,1)*qnorm(0.975)*se
pvalue=1-pchisq(z^2, 1)
result=cbind(diff, ci[1], ci[2], pvalue)
colnames(result)=c("difference in DOR", "CI.low", "CI.high", "p.value")
print(round(result, 4))
```

```
##      difference in DOR CI.low CI.high p.value
```



```
## [1,]          0.6026 0.1971  1.0081  0.0036
```

In this example, the mean duration of response within a time window $[0, 6.75]$ between two groups is statistically significantly different with a p-value of 0.004. The estimated difference is 0.603 (95% confidence interval: 0.197 to 1.008). This p-value can also be viewed as that for the global comparison between two PIBR curves over the interval of $[0, 6.75]$.

Inferences on Cumulative Response Rate (CRR)

This package also allow making inference on the cumulative response rates between two groups using competing risk model. The cumulative response rate at time t is defined as

$$CRR(t) = P(T_1 \leq t) = P\{T_3 \leq t, T_1 < T_2\}.$$

Note $CRR(t)$ is a non-decreasing function unlike PBIR. It can be estimated as the cumulative incidence function in a competing risk setting. “PBIR” package called the output from the R function “cuminc” to make inferences about CRR. Here is the R code to estimate and compare CRR from two treatment groups:

```
fit=CRR(t2PROGRESSION=t2progression, STATUS_PROGRESSION=delta_progression,
        t2RESPONSE=t2response,      STATUS_RESPONSE=delta_response,
        TRT=trt, time=c(1,2,3,4,5))

fit
```

```
## $result0
##   time      CRR      se   ci.low   ci.up
## 1    1 0.2100000 0.04106834 0.1406246 0.3015898
## 2    2 0.2500000 0.04374768 0.1742339 0.3449501
## 3    3 0.2900000 0.04600179 0.2086181 0.3875808
## 4    4 0.3136364 0.04756968 0.2285913 0.4133669
## 5    5 0.3136364 0.04756968 0.2285913 0.4133669
##
## $result1
##   time      CRR      se   ci.low   ci.up
## 1    1 0.4500000 0.05014355 0.3548564 0.5489473
## 2    2 0.5700000 0.05003967 0.4704605 0.6641845
## 3    3 0.6000000 0.04960436 0.5000926 0.6922288
## 4    4 0.6233333 0.04956191 0.5224822 0.7145227
## 5    5 0.6233333 0.04956191 0.5224822 0.7145227
##
## $pvalue
## [1] 6.039439e-06
```

The output includes the estimated CRR, its standard error and 95% confidence interval by treatment arm: result0 for TRT=0 vs result1 for TRT=1. The output also includes a p-value from comparing these two CRRs (Gray, 1988), which <0.001 in this example. One may also use the following R-code to plot the CRR by treatment arm

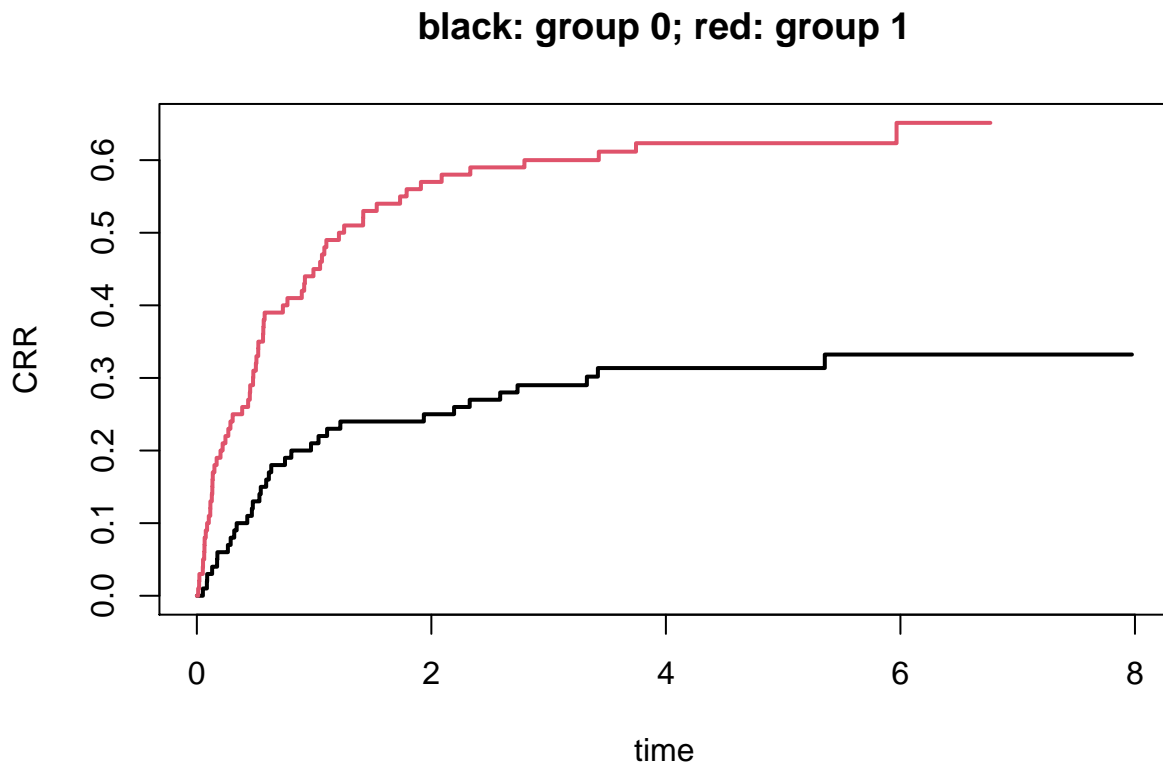
```
# Estimate the CRR over all the jump points of step functions
fit=CRR(t2PROGRESSION=t2progression,
        STATUS_PROGRESSION=delta_progression,
        t2RESPONSE=t2response,
        STATUS_RESPONSE=delta_response,
        TRT=trt)

# Plot the estimated PBIR by group
```

```

tt1=c(0, fit$result1$time)
crr1=c(0, fit$result1$CRR)
B1=length(tt1)
tt1=rep(tt1, rep(2, B1))[-1]
crr1=rep(crr1, rep(2, B1))[-(2*B1)]
tt0=c(0, fit$result0$time)
crr0=c(0, fit$result0$CRR)
B0=length(tt0)
tt0=rep(tt0, rep(2, B0))[-1]
crr0=rep(crr0, rep(2, B0))[-(2*B0)]
plot(range(c(fit$result1$time, fit$result0$time)),
      range(c(fit$result1$CRR, fit$result0$CRR)),
      xlab="time", ylab="CRR",
      main="black: group 0; red: group 1", type="n")
lines(tt0, crr0, col=1, lwd=2)
lines(tt1, crr1, col=2, lwd=2)

```



REFERENCE

Tsai W.Y., Luo X., Crowley J. (2017) The Probability of Being in Response Function and Its Applications. In: Matsui S., Crowley J. (eds) *Frontiers of Biostatistical Methods and Applications in Clinical Oncology*. Springer, Singapore. <doi: https://doi.org/10.1007/978-981-10-0126-0_10>.

Huang, B., Tian, L., McCaw, Z.R., Luo, X., Talukder, E., Rothenberg, M., Xie, W., Choueiri, T.K., Kim, D. H., & Wei, L. J. (2020). Analysis of response data for assessing treatment effects in comparative clinical

studies. *Annals of Internal Medicine* <doi: 10.7326/M20-0104>.

Huang, B., Tian, L., Talukder, E., Rothenberg, M., Kim, D. H., & Wei, L. J. (2018). Evaluating treatment effect based on duration of response for a comparative oncology study. *JAMA oncology*, 4(6), 874-876.

Gray, R.J. (1988) A class of K-sample tests for comparing the cumulative incidence of a competing risk, *Annals Of Statistics*, 16:1141-1154.