```
R code
> install.packages("BuyseTest", quiet = TRUE)
> library(BuyseTest)
```

```
R output
Loading required package: Rcpp
BuyseTest version 3.1.0
```

```
R code
> set.seed(10)
> data <- simBuyseTest(100, n.strata = 2)
> head(data)
```

```
R output
      id treatment  eventtime status toxicity      score strata
   <num>    <fctr>     <num>  <num>   <fctr>      <num> <fctr>
1:     1         C 0.17392093      1      yes -2.1250686      a
2:     2         C 0.16255166      0      yes  0.5211787      a
3:     3         C 0.08302502      1      yes -0.0464229      b
4:     4         C 0.22204972      0       no -1.1494717      b
5:     5         C 0.11669726      1       no  0.6293383      a
6:     6         C 0.11885540      1      yes -0.7264715      a
```

```
R code
> e.BT <- BuyseTest(treatment ~ tte(eventtime, status = status),
                    data = data)
```

```
R output
        Generalized Pairwise Comparisons

Settings
- 2 groups  : Control = C and Treatment = T
- 1 endpoint:
priority endpoint   type           operator             event
1        eventtime  time to event  higher is favorable  status (0 1)
- right-censored pairs: probabilistic score based on the survival curves

Point estimation and calculation of the iid decomposition

Estimation of the estimator's distribution
- method: moments of the U-statistic

Gather the results in a S4BuyseTest object
```

```
─────────────────────────────── R code ───────────────────────────────
> summary(e.BT)
```

```
────────────────────────────── R output ──────────────────────────────
      Generalized pairwise comparisons with 1 endpoint

- statistic      : net treatment benefit  (delta: endpoint specific, Delta: global)
- null hypothesis : Delta == 0
- confidence level: 0.95
- inference      : H-projection of order 1 after atanh transformation
- treatment groups: T (treatment) vs. C (control)
- censored pairs  : probabilistic score based on the survival curves
- results
endpoint total(%) favorable(%) unfavorable(%) neutral(%) uninf(%)  Delta CI [2.5% ; 97.5%] p.
eventtime     100        57.39          42.61          0        0 0.1479  [-0.0293;0.3161] 0.
```

```
─────────────────────────────── R code ───────────────────────────────
> confint(e.BT, statistic = "winRatio")
```

```
────────────────────────────── R output ──────────────────────────────
          estimate        se  lower.ci upper.ci null    p.value
eventtime 1.347081 0.2450411 0.9430953 1.924118    1 0.1014458
```

```
─────────────────────────────── R code ───────────────────────────────
> e.BThalf <- BuyseTest(treatment ~ tte(eventtime, status),
                   data = data, add.halfNeutral = TRUE, trace = FALSE)
> model.tables(e.BThalf, statistic = "favorable")
```

```
────────────────────────────── R output ──────────────────────────────
   endpoint total favorable unfavorable neutral uninf     Delta   lower.ci   upper.ci    p.value
1 eventtime   100  57.39388    42.60612       0     0 0.5739388 0.4852354 0.6581263 0.1019135
```

```
─────────────────────────────── R code ───────────────────────────────
> coef(e.BThalf, statistic = "winRatio")
```

```
────────────────────────────── R output ──────────────────────────────
[1] 1.347081
```

```
────────────────────── R code ──────────────────────
> e.MBT <- BuyseTest(treatment ~ tte(eventtime, status, threshold = 1) + bin(toxicity, operat
                    data = data, trace = 0)
> model.tables(e.MBT)
```

```
────────────────────── R output ──────────────────────
  endpoint threshold total favorable unfavorable neutral uninf   delta  Delta lower.ci upper
1 eventtime    1e+00 100.0      10.2        2.55    87.2     0  0.0768 0.0768 -0.00928     0.
3  toxicity    1e-12  87.2      18.8       24.72    43.7     0 -0.0590 0.0178 -0.13396     0.
```

```
────────────────────── R code ──────────────────────
plot(e.MBT)
```

```
────────────────────── R code ──────────────────────
plot(e.MBT, type = "racetrack")
```

```
────────────────────── R code ──────────────────────
> e.BTindiv <- BuyseTest(treatment ~ tte(eventtime, status = status),
                    data = data, keep.pairScore = TRUE)
> getPairScore(e.BTindiv)
```

```
────────────────────── R output ──────────────────────
       index.C index.T favorable unfavorable neutral      uninf weight
   1:        1     101 0.9192694  0.08073064       0 0.000000e+00      1
   2:        2     101 0.5695583  0.43044167       0 1.110223e-16      1
   3:        3     101 1.0000000  0.00000000       0 0.000000e+00      1
   4:        4     101 0.4969601  0.50303994       0 0.000000e+00      1
   5:        5     101 1.0000000  0.00000000       0 0.000000e+00      1
  ---
 9996:       96     200 0.2858328  0.71416716       0 0.000000e+00      1
 9997:       97     200 0.8120919  0.18790807       0 0.000000e+00      1
 9998:       98     200 0.6171644  0.38283561       0 0.000000e+00      1
 9999:       99     200 0.6171644  0.38283561       0 0.000000e+00      1
10000:      100     200 0.4596044  0.54039560       0 0.000000e+00      1
```

```
────────────────────── R code ──────────────────────
> eBT.perm <- BuyseTest(treatment ~ cont(score), data = data,
                    method.inference = "varexact permutation")
> model.tables(eBT.perm)
```

```
────────────────────── R output ──────────────────────
  endpoint total favorable unfavorable neutral uninf  Delta   p.value
1    score   100     53.67       46.33       0     0 0.0734 0.3698664
```

```
────────────────────── R code ──────────────────────
> wilcox.test(score ~ treatment, data = data, correct = FALSE)$p.value
```

```
────────────────────── R output ──────────────────────
0.3698664
```

```R code
> rbind(confint(e.BT, transformation = TRUE),
+       confint(e.BT, transformation = FALSE))
```

```R output
           estimate         se    lower.ci upper.ci null    p.value
eventtime  0.1478776 0.08897931 -0.02931684 0.3160612    0 0.10150573
eventtime1 0.1478776 0.08897931 -0.02651861 0.3222739    0 0.09652625
```

```R code
> NTB <- coef(e.BT)
> sigma.NTB <- sqrt(crossprod(getIid(e.BT)))
> sigmaTrans.NTB <- sigma.NTB/(1-NTB^2)
> c(estimate = NTB, se = sigmaTrans.NTB, p.value = 2*(1-pnorm(NTB/sigma.NTB)),
+   pTrans.value = 2*(1-pnorm(atanh(NTB)/sigmaTrans.NTB)))
```

```R output
   estimate         se     p.value pTrans.value
 0.14787764 0.09096860  0.09652625   0.10150573
```

```R code
BuyseTest.options(method.inference = "permutation", n.resampling = 1000,
                  statistic = "winRatio")
```

```R code
> data("prodige", package = "BuyseTest")
> head(prodige)
```

```R output
     id treatment      OS statusOS    PFS statusPFS toxicity    sex
<num>    <fctr>  <num>    <num>  <num>    <num>    <num> <fctr>
1:    1         C 0.0349        1 0.0349        0        1      F
2:    2         C 2.2790        0 2.2052        1        4      F
3:    3         C 0.2008        1 0.2008        0        1      M
4:    4         C 0.3418        1 0.3418        0        1      F
```

```R code
> e.BR <- BuyseTest(treatment ~ tte(OS, statusOS, threshold = 6)
+                   + cont(toxicity, operator = "<0", threshold = 2)
+                   + tte(OS, statusOS, threshold = 1)
+                   + cont(toxicity, operator = "<0"),
+                     data = prodige)
> plot(e.BR)
```

```R code
> summary(e.BR)
```

```
─────────────────────────── R code ───────────────────────────
> M.threshold <- cbind(OS_t6 = c(3:5,3:5),
                       toxicity_t2 = c(2,2,2,3,3,3),
                       OS_t2 = 1,
                       toxicity = 0)
> M.threshold
```

```
────────────────────────── R output ──────────────────────────
     OS_t6 toxicity_t2 OS_t2 toxicity
[1,]     3           2     1        0
[2,]     4           2     1        0
[3,]     5           2     1        0
[4,]     3           3     1        0
[5,]     4           3     1        0
[6,]     5           3     1        0
```

```
─────────────────────────── R code ───────────────────────────
> eBR.Se <- sensitivity(e.BR, band = TRUE,
                        threshold = M.threshold)
> plot(eBR.Se)
```

```
─────────────────────────── R code ───────────────────────────
> e.NTB_Gehan <- BuyseTest(treatment ~ tte(OS, statusOS), scoring.rule = "Gehan",
                           data = prodige, keep.pairScore = TRUE, trace = FALSE)
> getPairScore(e.NTB_Gehan)[1:2,]
```

```
────────────────────────── R output ──────────────────────────
   index.C index.T favorable unfavorable neutral uninf weight
1:       1     403         1           0       0     0      1
2:       2     403         0           0       0     1      1
```

```
─────────────────────────── R code ───────────────────────────
> e.NTB_Peron <- BuyseTest(treatment ~ tte(OS, statusOS), scoring.rule = "Peron",
                           data = prodige, keep.pairScore = TRUE, trace = FALSE)
> getPairScore(e.NTB_Peron)[1:2,]
```

```
────────────────────────── R output ──────────────────────────
   index.C index.T favorable unfavorable neutral        uninf weight
1:       1     403 1.0000000     0.00000       0 0.0000000000      1
2:       2     403 0.5286551     0.47068       0 0.0006648516      1
```

```
─────────────────────────── R code ───────────────────────────
> e.NTB_Latta <- BuyseTest(treatment ~ tte(OS, statusOS), scoring.rule = "Peron",
                           data = prodige, trace = FALSE,
                           model.tte = prodlim(Hist(OS, statusOS) ~ 1, data = prodige))
```

```
─────────────────────────── R code ───────────────────────────
> e.NTB_restricted <- BuyseTest(treatment ~ tte(OS, statusOS, restriction = 24),
                                scoring.rule = "Peron", data = prodige)
```

```
────────────────────── R code ──────────────────────
> simFCT <- function(n.C, n.T){
    df.C <- data.frame(id = paste0("C",1:n.C), group = 0,
                       tox = sample(1:6, n.C, replace=TRUE,
                                    prob = c(16.09, 15.42, 33.26, 26.18, 8.38, 0.67)/100),
                       time = rweibull(n.C, scale = 9.995655, shape = 1.28993),
                       event = 1)
    df.T <- data.frame(id = paste0("T",1:n.T), group = 1,
                       tox = sample(1:6, n.T, replace=TRUE,
                                    prob = c(8.21, 13.09, 31.29, 30.87, 12.05, 4.49)/100),
                       time = rweibull(n.T, scale = 13.16543, shape = 1.575269),
                       event = 1)
    return(rbind(df.C,df.T))
}
> set.seed(10)
> simFCT(2,2)
```

```
────────────────────── R output ──────────────────────
  id group tox      time event
1 C1     0   4  8.821945     1
2 C2     0   3  4.591318     1
3 T1     1   3 15.495787     1
4 T2     1   3 15.557655     1
```

```
────────────────────── R code ──────────────────────
> e.power <- powerBuyseTest(group ~ tte(time,event,threshold = 1)  + cont(tox, operator = "<(
                            sim = simFCT, sample.size = c(10,25,50),
                            n.rep = 100, seed = 10, cpus = 1)
> summary(e.power)
```

```
────────────────────── R output ──────────────────────
        Simulation study with Generalized pairwise comparison with 100 samples

- net benefit statistic (null hypothesis Delta=0)
endpoint threshold n.T n.C mean.estimate sd.estimate mean.se rejection.rate
     tox     1e-12  10  10        0.2156      0.2656  0.2468           0.13
                   25  25        0.2032      0.1677  0.1582            0.2
                   50  50        0.2015      0.1228  0.1121           0.43


n.T          : number of observations in the treatment group
n.C          : number of observations in the control group
mean.estimate: average estimate over simulations
sd.estimate  : standard deviation of the estimate over simulations
mean.se      : average estimated standard error of the estimate over simulations
rejection    : frequency of the rejection of the null hypothesis over simulations
(standard error: H-projection of order 1| p-value: after transformation)
```

```
──────────────── R code ────────────────
> e.n <- powerBuyseTest(group ~ tte(time,event, threshold = 1) + cont(tox, operator = "<0"),
                        sim = simFCT, power = 0.8,
                        n.rep = c(1000,10), seed = 10, trace = 2, cpus = 1)
> summary(e.n)
```

```
──────────────── R output ────────────────
        Sample size calculation with Generalized pairwise comparison
        for a power of 0.8 and type 1 error rate of 0.05

- estimated sample size (mean [min;max]): 126 [91;155] controls
126 [91;155] treated

- net benefit statistic (null hypothesis Delta=0)
endpoint threshold n.T n.C mean.estimate sd.estimate mean.se rejection.rate
     tox     1e-12 126 126        0.2049       0.069  0.0707          0.818


n.T           : number of observations in the treatment group
n.C           : number of observations in the control group
mean.estimate: average estimate over simulations
sd.estimate  : standard deviation of the estimate over simulations
mean.se      : average estimated standard error of the estimate over simulations
rejection    : frequency of the rejection of the null hypothesis over simulations
(standard error: H-projection of order 1| p-value: after transformation)
```

$$U - \Delta = \underbrace{\frac{1}{m}\sum_{i=1}^{m} h_E(i)}_{\text{Experimental group}} + \underbrace{\frac{1}{n}\sum_{j=1}^{n} h_C(j)}_{\text{Control group}} + \underbrace{\frac{1}{mn}\sum_{i=1}^{n}\sum_{j=1}^{m} h_{EC}(i,j)}_{\text{Second order term}}$$

$$\text{where for } i \in \{1,\ldots,m\}, h_E(i) = \mathbb{E}[\mathbb{1}_{Y_i^E > Y_j^C} - \mathbb{1}_{Y_j^C > Y_i^E} \mid Y_i^E] - \Delta$$

$$j \in \{1,\ldots,n\}, h_C(j) = \mathbb{E}[\mathbb{1}_{Y_i^E > Y_j^C} - \mathbb{1}_{Y_j^C > Y_i^E} \mid Y_j^C] - \Delta$$

$$\widehat{\sigma}_U \underbrace{\approx}_{\text{First order}} \frac{1}{m^2}\sum_{i=1}^{m} h_E^2(i) + \frac{1}{n^2}\sum_{j=1}^{n} h_C^2(j)$$

$$p^{\mathcal{P}} = \frac{1}{1+P}\left\{1 + \sum_{p=1}^{P} \mathbb{1}_{|\Delta^{\mathcal{P}(p)}| \geq |\Delta|}\right\}$$

$$\mathbb{P}\left[Y_i^E > Y_j^C | \widetilde{Y}_i^E, \Omega_i^E, \widetilde{Y}_j^C, \Omega_j^C\right] = \begin{cases} 0.75 \text{ for } i = i_1 \\ 1 \text{ for } i = i_2 \end{cases}$$

$$\left(\widetilde{Y}_{i_1}^E, \widetilde{Y}_{i_2}^E, \widetilde{Y}_j^C, \Omega_{i_1}^E, \Omega_{i_2}^E, \Omega_j^C\right) = (4.7, 6.1, 1.5, 1, 1, 0)$$

$$U_{ij} = \mathbb{P}\left[Y_i > Y_j + \tau | \widetilde{y}_i, \omega_i, \widetilde{y}_j, \omega_j\right] - \mathbb{P}\left[Y_j > Y_i + \tau | \widetilde{y}_i, \omega_i, \widetilde{y}_j, \omega_j\right]$$

(ω_i, ω_j)	$\tilde{y}_i - \tilde{y}_j \leq -\tau$	$|\tilde{y}_i - \tilde{y}_j| < \tau$	$\tilde{y}_i - \tilde{y}_j \geq \tau$
$(1, 1)$	-1	0	1
$(0, 1)$	$\frac{S(\tilde{y}_j+\tau)+S(\tilde{y}_j-\tau)}{S(\tilde{y}_i)} - 1$	$\frac{S(\tilde{y}_j+\tau)}{S(\tilde{y}_i)}$	1
$(1, 0)$	-1	$-\frac{S(\tilde{y}_i+\tau)}{S(\tilde{y}_j)}$	$1 - \frac{S(\tilde{y}_i+\tau)+S(\tilde{y}_i-\tau)}{S(\tilde{y}_j)}$
$(0, 0)$	$\ldots$	$\frac{\int_{\tilde{y}_i}^{\infty} S(t+\tau)dS(t) - \int_{\tilde{y}_j}^{\infty} S(t+\tau)dS(t)}{S(\tilde{y}_i)S(\tilde{y}_j)}$	$\ldots$

$$\mathbb{P}\left[Y^E > Y^C\right] - \mathbb{P}\left[Y^C > Y^E\right]$$

$$\mathbb{P}\left[Y^E > Y^C\right] + 0.5\mathbb{P}\left[Y^E = Y^C\right]$$

$$\frac{\mathbb{P}\left[Y^E > Y^C\right] + 0.5\mathbb{P}\left[Y^E = Y^C\right]}{\mathbb{P}\left[Y^C > Y^E\right] + 0.5\mathbb{P}\left[Y^C = Y^E\right]}$$

$$\frac{\mathbb{P}\left[Y^E > Y^C\right]}{\mathbb{P}\left[Y^C > Y^E\right]}$$

$$\delta_1 = \mathbb{P}\left[Y_1^E > Y_1^C + \tau_1\right] - \mathbb{P}\left[Y_1^C > Y_1^E + \tau_1\right]$$
$$\delta_2 = \mathbb{P}\left[Y_2^E > Y_2^C \,\middle|\, |Y_1^E - Y_1^C| \le \tau_1\right] - \mathbb{P}\left[Y_2^C > Y_2^E \,\middle|\, |Y_1^E - Y_1^C| \le \tau_1\right]$$

$$\Delta_1 = \delta_1$$
$$\Delta_2 = \delta_1 + \delta_2$$

$$\Delta_1 = \delta_1$$
$$\Delta_2 = \delta_1 + \delta_2$$

$$\delta_1 = \mathbb{P}\left[Y_1^E > Y_1^C + \tau_1\right] - \mathbb{P}\left[Y_1^C > Y_1^E + \tau_1\right]$$
$$\delta_2 = \mathbb{P}\left[Y_2^E > Y_2^C \,\middle|\, |Y_1^E - Y_1^C| \le \tau_1\right] - \mathbb{P}\left[Y_2^C > Y_2^E \,\middle|\, |Y_1^E - Y_1^C| \le \tau_1\right]$$