

【日本製薬工業協会シンポジウム】

生存時間解析の評価指標に関する最近の展開

—RMST (restricted mean survival time) を理解する—

3. RMSTの必要症例数計算



2018年6月13日

医薬品評価委員会 データサイエンス部会 タスクフォース4
生存時間解析チーム

大鵬薬品工業(株) 棚瀬 貴紀

1. RMSTの症例数設計の方法
2. SASマクロの紹介
3. 症例数設計の例
 - 指数分布の場合
 - 区分指数分布の場合
4. まとめ

1. RMSTの症例数設計の方法

➤ 連続型確率変数の2群間の差

$$N = (1 + f) \frac{(z_{\alpha/2} + z_{\beta})^2}{(\mu_1 - \mu_0)^2 / (\sigma_1^2 / f + \sigma_0^2)}$$

μ_g : 群 g の平均, σ_g^2 : 群 g の分散,

α : 両側有意水準, $1 - \beta$: 検出力,

z_{α} : 標準正規分布の上側 $100\alpha\%$ 点,

f : 割付比 (実薬群 / 対照薬群)

➤ RMSTの差も基本的に同じだが, 打ち切りを考慮することが必要

- 目的: 2群間のRMSTの差を検出すること
- Royston and Parmar (2013)を参照
- 実際の臨床試験では, 境界時間 τ 内の打ち切り発生により分散が増加
- シミュレーションを用いて打ち切りを想定した分散を計算し, 対応する必要症例数を計算

1. 症例数設計で想定する分布及び期間等を決める.
 - 各治療群の生存時間分布 (例えば、区分指数分布を仮定)
 - 境界時間 τ
 - 被験者の登録期間及び追跡期間
 - 単位時間あたりの打ち切り発生割合
 - 被験者の登録時期の分布
 - 症例数の割付比
 - 有意水準, 検出力

計算手順(2)

2. 各治療群の生存時間分布及び境界時間 τ の設定に基づき、各群のRMST $\mu_g = \mu_g(\tau)$, $X(\tau)$ の分散 $\sigma_g^2 = \sigma_g^2(\tau)$ を計算する。

症例数の割付比, 有意水準及び検出力の設定に基づき, 境界時間 τ 内での打ち切りが発生しないと仮定した, 必要症例数 N (両群)を計算する。

$$N = (1 + f) \frac{(z_{\alpha/2} + z_{\beta})^2}{(\mu_1 - \mu_0)^2 / (\sigma_1^2 / f + \sigma_0^2)}$$

α : 両側有意水準, $1 - \beta$: 検出力,

z_a : 標準正規分布の上側 $100a\%$ 点,

f : 割付比(実薬群/対照薬群)

計算手順(3)

3. 登録期間, 登録時期の分布, 追跡期間及び打ち切り発生を考慮したモンテカルロ・シミュレーションを行う.
 1. シミュレーションの各繰り返し s ($= 1, \dots, M$)について, RMSTの標準誤差 $\hat{\sigma}(\hat{\mu}_{g,s})$ を推定する.
 2. RMSTの分散の調整係数 $\phi_{g,s} = \sqrt{m}\hat{\sigma}(\hat{\mu}_{g,s})/\sigma_g$ から, RMSTの分散 $\sigma_{g,s}^2 = \phi_{g,s}^2\sigma_g^2$ を計算する (m :シミュレーションでの各群の症例数).
 3. 打ち切りを考慮した必要症例数 N_s を計算する.

$$N_s = (1 + f) \frac{(z_{\alpha/2} + z_{\beta})^2}{(\mu_1 - \mu_0)^2 / (\sigma_{1,s}^2 / f + \sigma_{0,s}^2)}$$

計算手順(4)

4. 打ち切りを考慮した必要症例数 N_S の平均及び標準誤差を計算する.

標準誤差は次式で求める. 標準誤差はシミュレーションの結果が安定しているかどうかを確認するために求める.

シミュレーションの標準誤差が実用上十分に小さいと考えられるシミュレーション回数 M を選ぶ.

$$SE(N_S) = \sqrt{(N_S \text{の標本分散})/M}$$

2. SASマクロの紹介

- タスクフォースにて、必要症例数を計算するSASマクロを作成
- SASマクロは報告書に掲載されています
- 他、RのSSRMSTパッケージによる計算も可能 (Horiguchi et al. (2017))

SASマクロの入力引数(1)

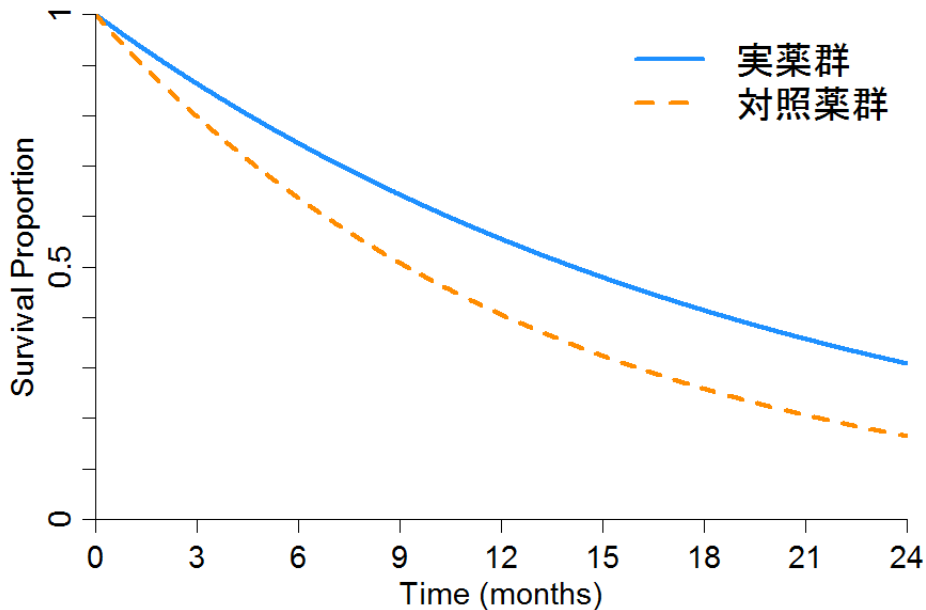
引数	説明
P1, P0	引数timeに対応する実薬群及び対照薬群の生存割合のベクトル
time	区分指数分布の区分時間
tau	境界時間 τ
enroll	被験者の登録期間
follow	被験者の追跡期間(被験者の登録が終了した時点から観察終了までの期間)
loss	単位時間あたりのランダム打ち切りの発生割合(単位は%), デフォルトは両群0%

SASマクロの入力引数(2)

引数	説明
r	登録時期の分布のパラメータ, デフォルトは1
ratio	症例数の割付比(実薬群:対照薬群), デフォルトは1:1
sig_level	両側有意水準 α , デフォルトは0.05
power	検出力 $1 - \beta$, デフォルトは0.8
seed	シミュレーションでの乱数のシード, デフォルトは123
ss_sim	シミュレーションでの各群の症例数 m , デフォルトは10,000
iteration	シミュレーションの繰り返し数 M , デフォルトは50

※ ss_sim及びiterationのデフォルト値は, Royston and Parmar (2013)の推奨による

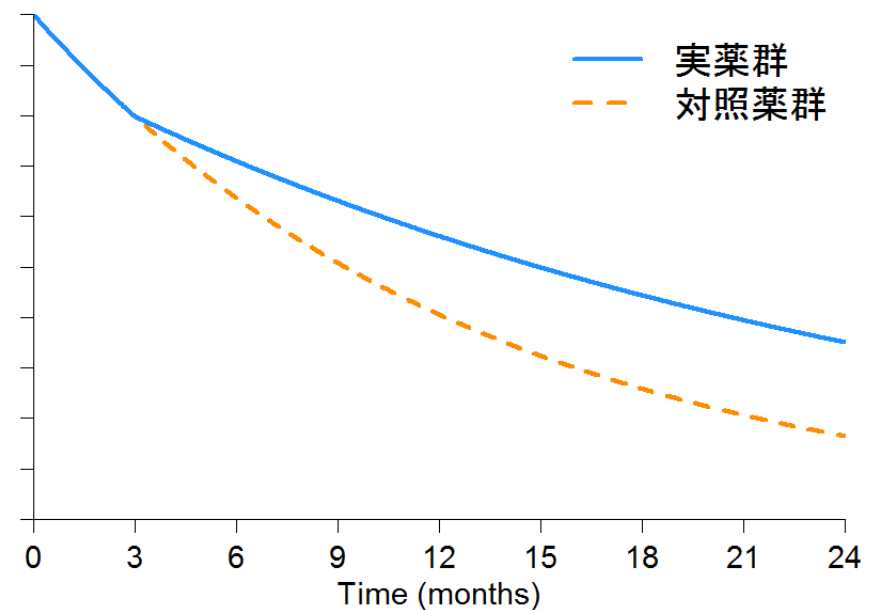
1. 指数分布



単純なモデルを仮定

- 区別が1つの区分指数分布

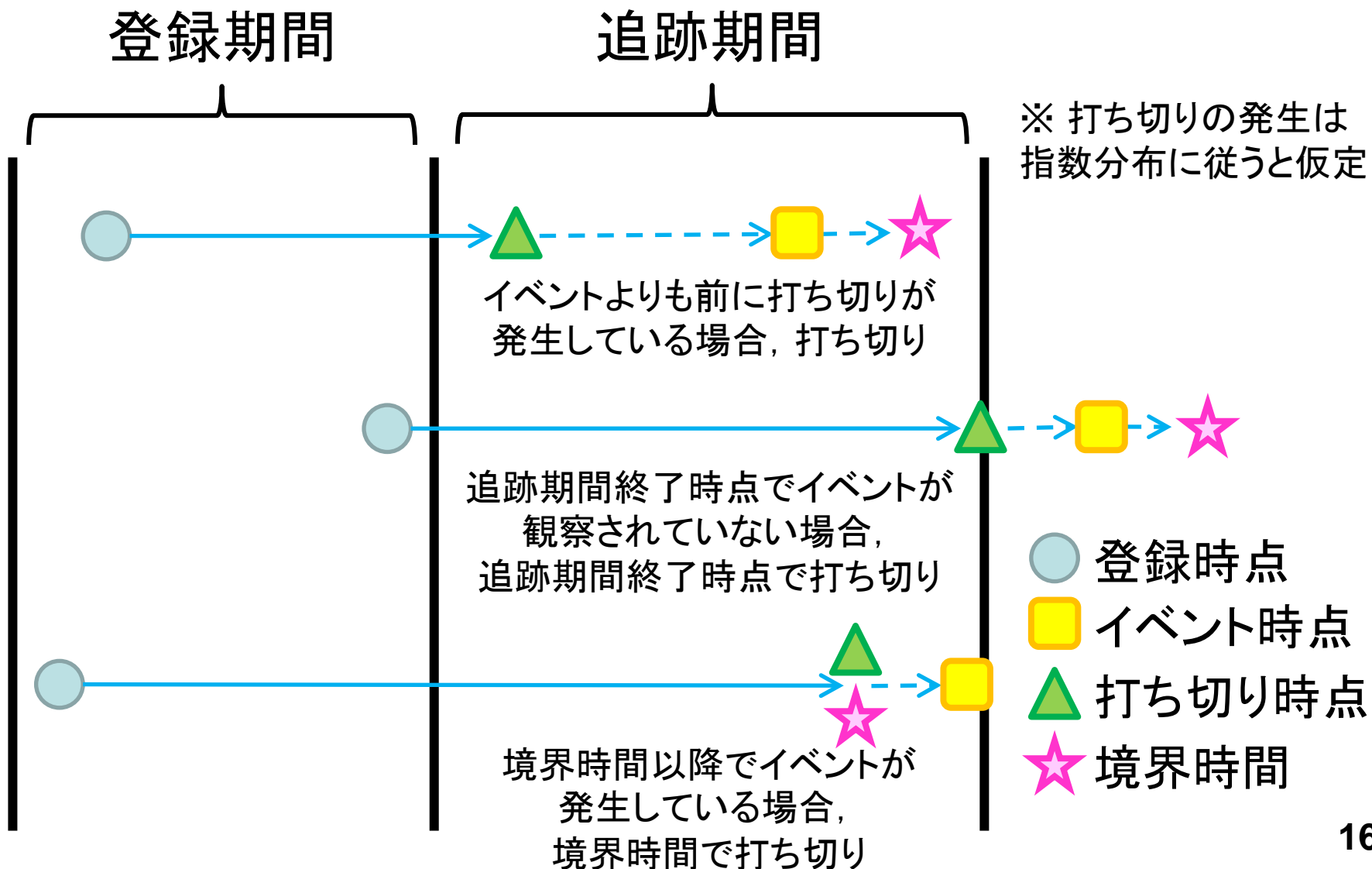
2. 区分指数分布



例: 遅発効果を仮定

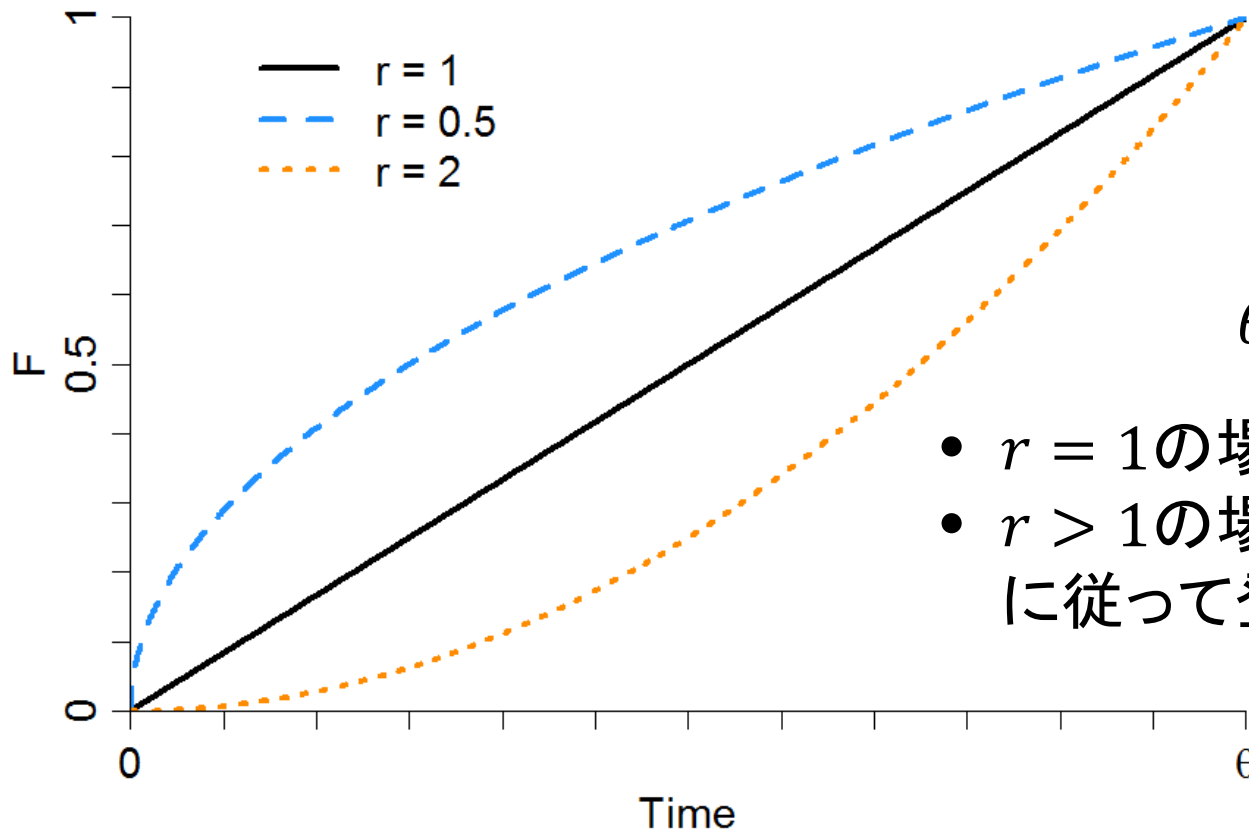
- 対照群は指数分布
- 実薬群はある時点まで対照群と同じ, それ以降で効果を発現

打ち切りのルール



登録時期の分布関数

$$F(t; \theta, r) = \left(\frac{t}{\theta}\right)^r, (r, \theta > 0)$$



θ : 登録期間

- $r = 1$ の場合, 一様分布に一致
- $r > 1$ の場合, 試験時期の進捗に従って登録スピードが増加

引数	説明
MEAN1	実薬群のRMST
MEAN0	対照薬群のRMST
VAR1	打ち切りを考慮しない実薬群の $X(\tau)$ の分散
VAR0	打ち切りを考慮しない対照薬群の $X(\tau)$ の分散
N	打ち切りを考慮しない必要症例数
MEAN_N	打ち切りを考慮した必要症例数のシミュレーション平均
SE_N	打ち切りを考慮した必要症例数のシミュレーション標準誤差

3. 症例数設計の例

➤ Overall survival

- ハザード比
0.73 (96% CI, 0.59–0.89)
- 24ヶ月RMST
13.0 vs 11.3 months
- 24ヶ月RMSTの差
1.7 months (95% CI, 0.4–3.1)

Borghaei et al. (2015)の
Figure 1A Overall Survival
を参照

Hasegawa et al. (2016)
Pak et al. (2017)

➤ 指数分布

■ 生存関数: $S(t) = \exp(-\lambda t)$

➤ $X(\tau)$ の平均 (RMST), 分散

$$\mu(\tau) = \frac{1 - \exp(-\lambda\tau)}{\lambda}$$

$$\sigma^2(\tau) = \frac{1 - 2\lambda\tau\exp(-\lambda\tau) - \exp(-2\lambda\tau)}{\lambda^2}$$

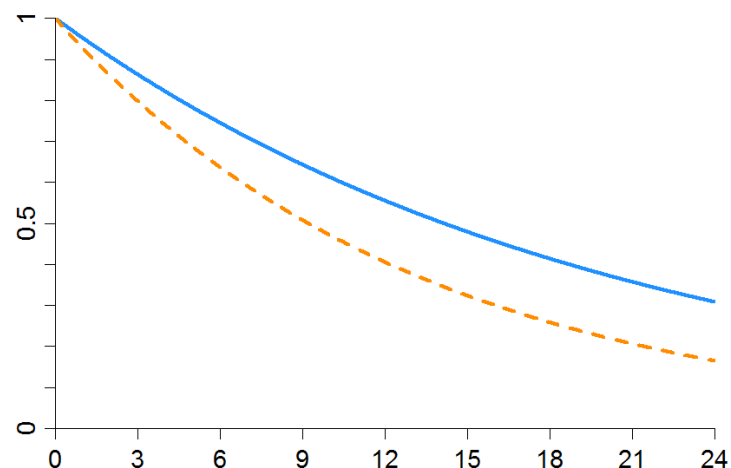
➤ RMSTの群間差3.0ヶ月

- 14.1ヶ月 vs 11.1ヶ月, 境界時間 $\tau = 24$
- 指数分布の下でのRMST

$$\mu_g(\tau) = \frac{1 - \exp(-\lambda_g \tau)}{\lambda_g}$$

Newton-Raphson法を用いて,
(SASコードは報告書に掲載)

実薬群 $\lambda_1 = 0.049088$, 対照薬群 $\lambda_0 = 0.075308$



24ヶ月生存割合は,

実薬群: $\exp(-0.049088 * 24) = 0.3078597$

対照薬群: $\exp(-0.075308 * 24) = 0.1640817$

➤ 区分指数分布

- 既知の時点 $t_{(1)}, \dots, t_{(k)}$, 境界時間 $\tau > t_{(k)}$

- 時間の区分

$$(t_{(0)} = 0, t_{(1)}], (t_{(1)}, t_{(2)}], (t_{(2)}, t_{(3)}], \dots, (t_{(k)}, t_{(k+1)} = \infty]$$

➤ 時間の区分幅

$$\delta_{j+1} = \begin{cases} t_{(j+1)} - t_{(j)}, & j < k \\ \tau - t_{(k)}, & j = k \end{cases}$$

➤ 生存関数

$$S_{j+1}(t) = \exp(-H_j) \exp(-h_{j+1}(t - t_{(j)}))$$

- h_{j+1} : 区間 $(t_{(k)}, t_{(k+1)})$ の指数ハザード

- $H_j = \sum_i h_i \delta_i$: 時間 $t_{(j)}$ までの累積指数ハザード

区分指数分布を仮定したRMST

➤ $X(\tau)$ の平均 (RMST), 分散

$$\mu(\tau) = \sum_{j=0}^k \exp(-H_j) B_{j+1}$$

$$\sigma^2(\tau) = 2 \sum_{j=0}^k \exp(-H_j) (A_{j+1} + t_j B_{j+1}) - \{\mu(\tau)\}^2$$

ただし,

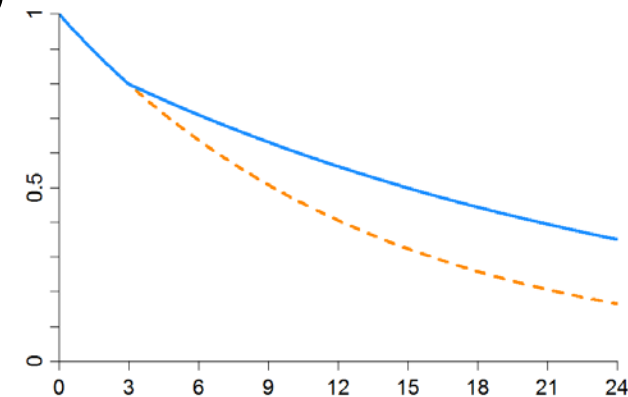
$$A_{j+1} = \frac{1}{h_{j+1}^2} \{1 - (1 + h_{j+1} \delta_{j+1}) \exp(-h_{j+1} \delta_{j+1})\},$$

$$B_{j+1} = \frac{1 - \exp(-h_{j+1} \delta_{j+1})}{h_{j+1}}$$

➤ RMSTの群間差3.0ヶ月

- 14.1ヶ月 vs 11.1ヶ月, 境界時間 $\tau = 24$
- 対照薬群: 指数分布 ($\lambda_0 = 0.075308$)
- 実薬群: 0~3ヶ月は対照群と同じ
3ヶ月以降で治療効果に差

$$\mu(\tau) = \sum_{j=0}^k \exp(-H_j) B_{j+1}$$



Newton-Raphson法を用いて(SASコードは報告書に掲載),
3~24ヶ月の指数ハザードを計算すると, $h_2 = 0.039219$

24ヶ月生存割合は,

実薬群: $\exp(-0.075308 * 3) * \exp(-0.039219 * 21) = 0.3501047$

対照薬群: $\exp(-0.075308 * 24) = 0.1640817$

各シナリオの設定 時間単位は全て月

設定	シナリオ1	シナリオ2	シナリオ3	シナリオ4
RMST	14.1 vs 11.1			
境界時間	24			
登録期間	11	18	11	11
追跡期間	15	8	15	15
ランダム打ち切り 発生割合	0	0	1 %	0
登録時期の 分布パラメータ	1	1	1	2
割付比	1:1			
両側有意水準	5 %			
検出力	90 %			

➤ シナリオ1

```
%SS_RMST(  
P1 = 0.3078597, P0 = 0.1640817,  
time = 24, tau = 24, power = 0.9,  
enroll = 11, follow = 15);
```

➤ シナリオ2

```
%SS_RMST(  
P1 = 0.3078597, P0 = 0.1640817,  
time = 24, tau = 24, power = 0.9,  
enroll = 18, follow = 8);
```

➤ シナリオ3

```
%SS_RMST(  
P1 = 0.3078597, P0 = 0.1640817,  
time = 24, tau = 24, power = 0.9,  
enroll = 11, follow = 15, loss = %str(0.01, 0.01)) ;
```

➤ シナリオ4

```
%SS_RMST(  
P1 = 0.3078597, P0 = 0.1640817,  
time = 24, tau = 24, power = 0.9,  
enroll = 18, follow = 8, r = 2) ;
```

必要症例数の計算結果 指数分布の場合

	シナリオ1	シナリオ2	シナリオ3	シナリオ4
RMST	14.1 vs 11.1			
打ち切りを考慮しない $X(\tau)$ の分散	74.6325 vs 66.9967			
打ち切りを考慮しない 必要症例数 N	332			
打ち切りを考慮した必 要症例数のシミュレー ション平均(上段), 標準誤差(下段)	336 0.21244	366 0.27073	358 0.31981	346 0.22079

シミュレーション回数50回で、標準誤差は必要症例数に対して十分に小さい。

※ 計算時間は、1シナリオあたり30秒未満

(Windows 7 Pro 32bit, Core i5-2520M 2.50GHz, RAM 4GB)

➤ シナリオ1

```
%SS_RMST(  
P1 = %str(0.7977788, 0.3501062),  
P0 = %str(0.7977788, 0.1640817),  
time = %str(3, 24), tau = 24, power = 0.9,  
enroll = 11, follow = 15) ;
```

➤ シナリオ2

```
%SS_RMST(  
P1 = %str(0.7977788, 0.3501062),  
P0 = %str(0.7977788, 0.1640817),  
time = %str(3, 24), tau = 24, power = 0.9,  
enroll = 18, follow = 8) ;
```

➤ シナリオ3

```
%SS_RMST(  
P1 = %str(0.7977788, 0.3501062),  
P0 = %str(0.7977788, 0.1640817),  
time = %str(3, 24), tau = 24, power = 0.9,  
enroll = 11, follow = 15, loss = %str(0.01, 0.01)) ;
```

➤ シナリオ4

```
%SS_RMST(  
P1 = %str(0.7977788, 0.3501062),  
P0 = %str(0.7977788, 0.1640817),  
time = %str(3, 24), tau = 24, power = 0.9,  
enroll = 18, follow = 8, r = 2) ;
```

必要症例数の計算結果 区分指数分布の場合

	シナリオ1	シナリオ2	シナリオ3	シナリオ4
RMST	14.1 vs 11.1			
打ち切りを考慮しない $X(\tau)$ の分散	84.6029 vs 66.9967			
打ち切りを考慮しない 必要症例数 N	354			
打ち切りを考慮した必 要症例数のシミュレー ション平均(上段), 標準誤差(下段)	360 0.20846	388 0.28428	380 0.35165	368 0.2336

シミュレーション回数50回で、標準誤差は必要症例数に対して十分に小さい。

※ 計算時間は、1シナリオあたり30秒未満

(Windows 7 Pro 32bit, Core i5-2520M 2.50GHz, RAM 4GB)

- RMSTの群間差を検出するための症例数設計の方法を紹介した
- 打ち切りに関連する仮定(登録・追跡期間, 登録時期の分布, ランダム打ち切り)の設定によって, 必要症例数は変わる
- 仮定と現実には差があった場合でも安定して結論を得られるような, 症例数設計が重要

- Royston P, Parmar MKB. Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. *BMC Med Res Methodol*. 2013;13:152.
- Horiguchi M, Uno H. CRAN – Package SSRMST. <https://cran.r-project.org/web/packages/SSRMST/index.html>
- Borghaei H et al. Nivolumab versus docetaxel in advanced nonsquamous non–small-cell lung cancer. *N Eng J Med* 2015; 373:1627–1639.
- Hasegawa T et al. Nivolumab in nonsquamous non–small-cell lung cancer [Letter to the editor]. *N Eng J Med* 2016; 374:492–494.
- Pak K, Uno H, Kim DH, Tian L, Kane RC, Takeuchi M, et al. Interpretability of cancer clinical trial result using restricted mean survival time as an alternative to the hazard ratio. *JAMA Oncol*, 2017;3:1692–1696.
- Fine GD. Consequences of delayed treatment effects on analysis of time-to-event endpoints. *Drug Inf J*. 2007;41:535–539.