

SASによる生存時間解析の実務

張 方紅
グラクソ・スミスクライン(株)バイオメディカルデータサイエンス部

Practice of Survival Analysis sing SAS

Fanghong Zhang
Biomedical Data Science Department,
GlaxoSmithKline K.K.

要旨:

SASによる生存時間解析の実務経験を共有する。
データの要約・検定・推定の前提に注意し、群間差
の検定とハザード比の信頼区間に焦点を与える。複
数に存在するプログラミング方法の間の関係と使い
勝手を紹介する。

キーワード: 生存時間解析, LIFETESTプロシジャ, ロ
グランク検定, ハザード比

動機づけ:実務経験を共有

生存時間解析に報告が多い

- 概念・本質
 - 浜田(2011), 生存時間解析再入門 「生存時間解析のミステリーをひも解く」
- 学術
 - 平井・吉田・田崎(2009), 2つの生存時間の差に対する同時信頼区間
- 実務レベルの経験

発表内容

- 仮想な臨床試験
- SASプログラミング
- 感想
- まとめ

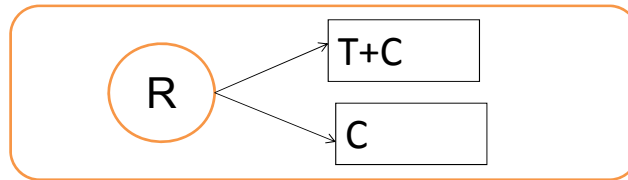
仮想な臨床試験

- 主要目的: 治験薬群であるT+C併用療法と対照薬群であるC単独療法とで全生存期間 (OS Overall Survival)を比較する
- 主要評価項目: 全生存期間

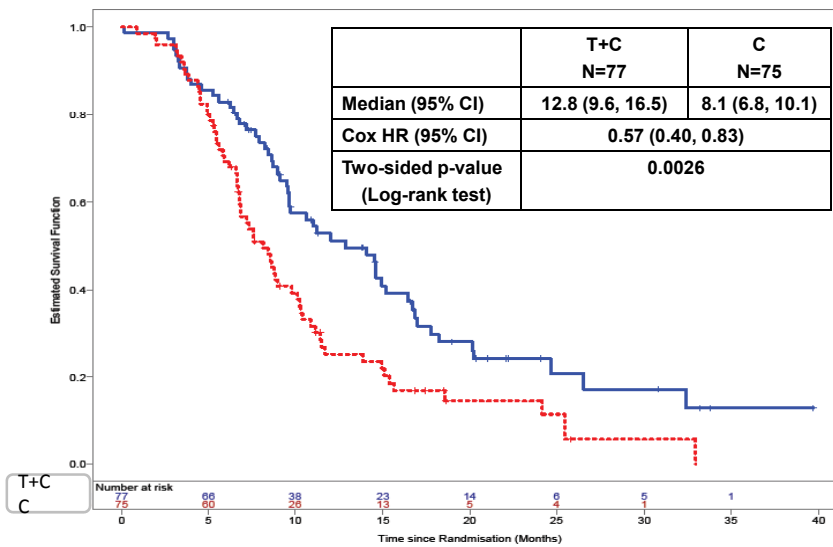
解析方法

1. 累積生存曲線, 生存期間中央値の推定は Kaplan-Meier法を用いて行い、Greenwoodの公式を用いて95%信頼区間を求める
2. 群間比較には施設以外の割付調整因子を層とした層別ログランク検定を用いる
3. 治療効果の推定値として、Coxの比例ハザードモデルを用いて群間の治療効果のハザード比とその95%信頼区間を求める

ASCO (American Society of Clinical Oncology)



Overall survival



レポート: 結果

- 主要評価項目であるOSの中央値はT+C併用群で12.8ヵ月、C単独群で8.1ヵ月 (HR=0.57, 95%CI: 0.40-0.83, p=0.0026) と有意な延長を示した.

HR: Hazard Ratio

CI: Confidence Interval

解析段階と前提

解析段階	内容	前提
分布の記述 要約	生存関数をKaplan-Meier法で推定し、プロットする。推定された生存曲線から中央値を求める	ノンパラメトリック法であり、前提が 必要なし
検定	群間で生存関数の違いがあるかを検定する	ノンパラメトリック法、生存関数が群間で交差していない
推定	ハザード比を推定する	セミパラメトリック法、比例ハザード性

ログランク検定の帰無仮説

- 両群のOSの中央値が等しいという仮説に対する検定→×
- 両群のハザード比=1という仮説に対する検定→×
- 両群の生存関数が等しいという仮説に対する検定→○

中央値・ハザード比・P値

	T+C N=58	C N=57
Median (95% CI)	5.6 (5.3, 7.7)	4.4 (2.8, 5.7)
HR (95% CI) (Cox)	0.61 (0.37, 1.00)	
Two-sided p-value (Log-rank test)	0.0315	

CI of HRとP値の結果と一致しない

- 前提が異なる
- 手法が対応していない: Wald vs Score

ここまでの結論

- ログランク検定は両群の生存関数が等しいという仮説に対する検定
- 中央値・ハザード比・ログランク検定はそれぞれ前提が異なる, 方法が対応していない場合が多い

ログランク検定: 概念・記号

ログランク(log-rank)検定は, 生存曲線の差を、ある時点ではなく観察期間全体を通じて比較し検定できる方法です.

$$\chi^2 = \frac{(O - E)^2}{V}$$

O-E:スコア統計量

V: O-Eの分散

O: 治験薬群における観察死亡数, E: 治験薬群における期待死亡数

ログランク検定: SASプロシジャ

1	PROC LIFETEST; STRATA 治療群変数; RUN;
2	PROC LIFETEST; TEST 治療群変数; RUN;
3	PROC PHREG; MODEL 時間変数*打ち切り変数=治療群変数/TIES=; RUN;

3つの方法で本質的に同等な結果
違いはタイの取り扱い

大橋・浜田 (1995), p.54

ログランク検定: SASプロシジャ

PROC LIFETEST; STRATA 治療群変数; RUN;	PROC PHREG; MODEL 時間変数*打ち切り変数=治療群変数/ TIES=DISCRETE; RUN;
PROC LIFETEST; TEST 治療群変数; RUN;	PROC PHREG; MODEL 時間変数*打ち切り変数=治療群変数/ TIES=BRESLOW; /*デフォルト*/ RUN;

上の表の左・右の方法はそれぞれ結果が一致
ただし, PHREGプロシジャのScore検定を使用

Collett (1994), p.332

ログランク検定: STRATA文を選ぶ

1	PROC LIFETEST; STRATA 治療群変数; RUN;
2	PROC LIFETEST; TEST 治療群変数; RUN;
3	PROC PHREG; MODEL 時間変数*打ち切り変数=治療群変数/TIES=; RUN;

STRATA文を選ぶ理由

1. 生存関数を計算
2. スコア統計量(O-E)を出力

層別ログランク検定: SASプロシジャ

1	PROC LIFETEST; STRATA 層別変数/GROUP=治療群変数; RUN;
2	PROC LIFETEST; STRATA 層別変数; TEST 治療群変数; RUN;
3	PROC PHREG; MODEL 時間変数*打ち切り変数=治療群変数/TIES=; STRATA 層別変数; RUN;

3つの方法で本質的に同等な結果

SAS 9.1 マニュアル

層別ログランク検定: GROUPオプション

```
1 PROC LIFETEST;  
  STRATA 層別変数/GROUP=治療群変数;  
  RUN;
```

```
2 PROC LIFETEST;  
  STRATA 治療群変数;  
  BY 層別変数;  
  RUN;
```

大橋・浜田(1995), p99-100

2つの方法で同じ結果
SAS 9.1 GROUPオプションが便利になった

ログランク検定: 片側

1. ログランク検定は両側しかない
2. 解決方法: ログランク検定に対応する漸近正規分布に従うZ統計量を利用

$$Z = \frac{O - E}{\sqrt{V}}$$

O-E: スコア統計量

V: O-Eの分散

ログランク検定：片側

```
ods output HomStats=stats /* O-Eスコア統計量*/
LogHomCov=cov; /*(V: O-Eの分散)*/
proc lifetest;
strata 治療群変数;
run;
```

ログランク検定：まとめ

両側	層なし	PROC LIFETEST; STRATA 治療群変数; RUN;
	層あり	PROC LIFETEST; STRATA 層別変数/GROUP=治療群変数; RUN;
片側	スコアO-Eとその分散Vをデータセットに落とし、Zを計算 ods output HomStats=stats LogHomCov=cov;	
	層なし	PROC LIFETEST; STRATA 治療群変数; RUN;
	層あり	PROC LIFETEST; STRATA 層別変数/GROUP=治療群変数; RUN;

ハザード比

1. Cox法
2. Peto法
3. Pike法
4. Mantel-Haenszel 法

ハザード比: 比例ハザード性を前提

$$\lambda_T(t) = \exp(\beta)\lambda_C(t)$$

λ_T : 治験薬群におけるハザード

λ_C : 対照薬群におけるハザード

比例ハザード性が成り立っていれば, 群間のリスクの違いを1つのパラメータに要約することが可能となる.

Cox法 : 部分尤度

$$L(\beta) = \prod_{j=1}^r \frac{\exp(\beta x_{(j)})}{\sum_{l=1}^{n_j} \exp(\beta x_l)}$$

Collett p.107, (3.28), $x_{(j)}=1$ 治験薬群, 0 対照薬群

O : 治験薬群における総死亡数, n_{1j} と n_{2j} はリスク集合の大きさ

$$\begin{aligned} \log L(\beta) &= \sum_{j=1}^r \beta x_{(j)} - \sum_{j=1}^r \log\left(\sum_{l=1}^{n_j} \exp(\beta x_l)\right) \\ &= \beta O - \sum_{j=1}^r \log(n_{1j} + n_{2j} \exp(\beta)) \end{aligned}$$

Cox法

$$\begin{aligned} u(\beta) &= \frac{\partial \log L(\beta)}{\partial \beta} = O - \sum_{j=1}^r \frac{n_{2j} \exp(\beta)}{n_{1j} + n_{2j} \exp(\beta)} \\ &= O - E | H_1 \end{aligned}$$

$E|H_1$: 治験薬群における対立仮説の下での期待死亡数

Cox法 : SAS

 $\hat{\beta}_{ML}$: 最尤推定値

$$u(\hat{\beta}_{ML}) = 0, O = E | H_1$$

$$CI \text{ of } \beta = \hat{\beta}_{ML} \pm 1.96SE(\hat{\beta}_{ML})$$

$$CI \text{ of HR} = \exp(\hat{\beta}_{ML} \pm 1.96SE(\hat{\beta}_{ML}))$$

```
PROC PHREG;  
MODEL 時間変数*打ち切り変数=治療群変数/RL;  
RUN;
```

Peto法

 β_1 : β の最尤推定の one step 近似

$$\beta_1 = \frac{O - E}{V}$$

大橋・浜田 (1995), p.91-95

O-E: スコア統計量

V: O-E の分散

O: 治験薬群における観察死亡数,

E: 治験薬群における期待死亡数

Peto法 : SAS

$$\beta_1 = \frac{O-E}{V} \quad SE(\beta_1) = \frac{1}{\sqrt{V}}$$

$$CI \text{ of HR} = \exp\left(\frac{O-E}{V} \pm \frac{1.96}{\sqrt{V}}\right)$$

```
ods output HomStats=stats LogHomCov=cov;  
PROC LIFETEST;  
RUN;
```

Pike法

$$\tilde{\beta} = \log\left(\frac{O_T/E_T}{O_C/E_C}\right) \quad SE(\tilde{\beta}) = \sqrt{\frac{1}{E_T} + \frac{1}{E_C}}$$

$$CI \text{ of HR} = \exp\left(\log\frac{O_T/E_T}{O_C/E_C} \pm 1.96\sqrt{\frac{1}{E_T} + \frac{1}{E_C}}\right)$$

Pike法 : SAS

$$CI \text{ of HR} = \exp \left(\log \frac{O_T / E_T}{O_C / E_C} \pm 1.96 \sqrt{\frac{1}{E_T} + \frac{1}{E_C}} \right)$$

計算対象	SASプログラミング
O	ods output censoredsummary=events; PROC LIFETEST; RUN;
O-E	ods output HomStats=stats; PROC LIFETEST; RUN;
E	$E=O-(O-E)$

感想

- ログランク検定(層別ありなし) : SAS 9.1以降は直接に出力でき、便利になった.
- ハザード比 : Cox法しか直接に出力できなく、不便のまま
- ハザード比に対して、方法論の研究と比較も不十分

まとめ

- 中央値・ハザード比・ログランク検定のP値はそれぞれ前提が異なる, 対応していない
- LIFETESTプロシジャSTRATA文は使い勝手がよい
- ハザード比に対する方法論の比較研究が必要

参考文献 (1/2)

- 大橋靖雄・浜田知久馬 (1995), 生存時間解析 SASによる生物統計. 東京大学出版会.
- 浜田知久馬 (2011), 生存時間解析再入門 「生存時間解析のミステリーをひも解く」. 日本SASユーザー会論文集, 3-43.
- 平井健太・吉田祐樹・田崎武信 (2009), 2つの生存時間の差に対する同時信頼区間. 日本SASユーザー会論文集, 127-151.
- JCOG (2008) JCOGプロトコールマニュアル version 2.0. http://www.jcog.jp/doctor/tool/14_20.pdf (2012.6アクセス)
- ASCO 2012 演題速報 <http://www.gi-cancer.net/gi/asco/2012/report/LBA3500/> (2012.6アクセス)

参考文献 (2/2)

- Berry G., Kitchin R. M. and Mock P. A. (1991): A comparison of two simple hazard ratio estimators based on the logrank test. *Statistics in Medicine*, Vol. 10, 749-755
- Collett D. (1994), *Modelling Survival Data in Medical Research*. Chapman & Hall: London.
- Sato T.(1992) Letters to the editor, *Statistics in Medicine*, p. 847-848
- SAS/STAT® 9.1 User's Guide The LIFETEST Procedure